

Children Learn to Read and Write Chinese Analytically

**Lily Chan
Child Development and Learning
Institute of Education, University of London**

Supervisors : Professor Terezinha Nunes

Ms. Angela Hobsbaum

A thesis submitted for the degree of Doctor of Philosophy

May, 1996



Abstract

Recent progress in psycholinguistic research on written Chinese allows us to develop a new approach to investigate the Chinese reading acquisition process. We hypothesized that Chinese children, much like children learning an alphabetic script, do not simply learn written words by rote. As they are taught words to be learned by rote, they develop an implicit understanding of the formal and functional characteristics of written Chinese. The formal characteristics refer to the graphic structure and the positioning of the stroke-patterns, and the functional characteristics refer to the semantic and phonological information conveyed in the stroke-patterns. The studies reported were designed to investigate the nature of children's learning of written Chinese. In two series of studies, a total of 236 children from Hong Kong, aged four to nine, created and decoded novel Chinese compound words. Results showed that young Chinese children attended to both the formal and functional constraints in reading and writing tasks. In the judging task, 4-year-olds were able to identify the type of orthographic elements - the stroke-patterns, but they could not place them in legitimate positions. The 6-years-olds were able to refer both to the position and the correct type of orthographic elements in differentiating pseudowords from nonwords. In the writing and reading tasks, four and five-year-olds were unable to utilize the semantic radicals to represent meaning, nor could they use the phonological components for pronunciation; six-year-olds could use the semantic radicals to represent meaning and only nine-year-olds could both use semantic radicals correctly and systematically referred to the phonological components for pronunciation. A significant age difference was found in all the experiments. The studies provide strong evidence that learning compound words in Chinese is not a simple matter of memorizing but involves the understanding of formal and functional constraints in the script. A possible application of these findings lies in the new direction offered for reading instruction where the non-generative, rote view of learning to read and write in Chinese can be safely abandoned.

ACKNOWLEDGMENTS

I was told when I was considering doing a Ph.D. that it is a long and lonely journey. I am glad that it is not the case for me because my supervisors, my family and my friends have made this journey a lot easier and pleasant.

I would like to express my most sincere thanks to Professor Terezinha Nunes. She is a superb supervisor. She is kind, considerate, inspiring and intelligent. She understands her students and always offers just the right kind of stretch to push her students to progress further. To me, she is Socrates, as she nourishes my mind and guides me to enter the fascinating world of knowledge. I would also like to thank Ms Angela Hobsbaum for her comments on my work and encouragement.

My family has been very supportive. They look after me when I am in Hong Kong, and take care of my flat when I am in London. Although my parents do not understand what a thesis is, they are always there to share my joy and distress.

I do not think I can finish my study without the valuable assistance from my friends. Maria, Hester, and Josephine helped me to collect the data for the pilot study. Maria inspired my interest in children's writing and all these years has been encouraging me to persist in my work. Samuel made a series of interesting pictures for me and they were very useful in stimulating children to read and write in this study. Gwenda sacrificed her Easter Holiday to read my final draft, and Simon helped me to do the final editing of tables and figures.

It is a gift from God that I can meet the fellow research students, Milita, Miriam and Ursula in spite of my short stay in London every year. From time to time, we spent hours for tea, to discuss our work and to share our doubts. When I look back, those are the most memorable moments in my Ph.D. pursuit. What joy it is to be able to build up acquaintances even at this critical period of our life.

I would like to thank the principals and teachers of Caritas Kindergarten, TWGHS Kwan Kai Ming Primary School and TWGHS Sin Chu Wan Primary School for allowing me to conduct my study at short notices. Many teachers provided every possible assistance to allow me to conduct the experiments smoothly. They also shared with me valuable experiences in the teaching of Chinese. Last, but not the least, I would like to thank the children who participated in this study. They were very co-operative, and shared with me their knowledge of written Chinese. I enjoyed every minute of the process of the data collection because of their active involvement.

TABLE OF CONTENTS

ABSTRACT	2
ACKNOWLEDGMENTS	3
TABLE OF CONTENTS	4
LIST OF TABLES	11
LIST OF FIGURES	13
 CHAPTER 1	
INTRODUCTION	14
 CHAPTER 2	
THE CHINESE LANGUAGE	20
2.1 Introduction	20
2.2 Spoken Chinese	20
2.3 Written Chinese	22
2.3.1 Six rules for word formation	22
2.3.2 Orthographic units of written Chinese	25
2.3.2.1 A stroke	25
2.3.2.2 A stroke-pattern	26
2.3.2.3 A character	29
2.3.2.4 A word	31
2.3.2.5 A sentence	32
2.3.3 Orthographic constituents of an ideophonic compound word	32
2.3.3.1 The semantic radical	33
2.3.3.2 The phonological component	36
2.3.3.2.1 The Orthography-Phonology Correspondence Rules	38
2.3.3.2.2 Phonological regularity of the Chinese orthography	39
2.3.4 Summary	41
 CHAPTER 3	
SKILLS OF FLUENT CHINESE READERS	
3.1 Introduction	43
3.2 Methodological issues	44
3.3 Phonological aspects of reading Chinese	46
3.3.1 Holistic approach	46
3.3.2 Analytical approach	52

3.4 Orthographic aspects of reading Chinese	56
3.4.1 Holistic approach	57
3.4.2 Analytical approach	60
3.5 Semantic aspects of reading Chinese	69
3.5.1 Holistic approach	69
3.5.2 Analytical approach	69
3.6 Conclusion on skills of fluent Chinese readers	74
3.7 Using analytical approach to investigate the processing of orthographic, semantic and phonological information in reading Chinese	75
3.7.1 Flores d'Arcais' (1992, 1995) studies	75
3.7.2 Chen's (1993, 1995) studies	77
3.7.2.1 The orthographic unit of Chinese words - the smallest stroke-pattern	78
3.7.2.2 Two functionally distinct orthographic constituents of compound words : the semantic radical and the phonological component	80
3.7.2.3 Orthographic-phonology correspondence rules : Derivation and Analogy	84
3.7.2.4 Conclusion on Chen's (1993, 1995) studies	86
CHAPTER 4	
CHILDREN LEARNING TO READ CHINESE	
4.1 Introduction	88
4.2 The pedagogic perspective on teaching written Chinese	88
4.2.1 Reading difficulties of Chinese children	88
4.2.2 Methods of teaching Chinese	90
4.3 Literacy development at pre-reading stage	91
4.3.1 Orthographic aspects of written Chinese	92
4.3.1.1 Developmental stages in understanding the orthographic structure of written Chinese	92
4.3.1.2 Developing global visual skills in reading Chinese	96
4.3.2 Phonological aspects of written Chinese	98
4.3.2.1 Developing one-to-one correspondence between a syllable and a character	98
4.3.2.2 Developing global phonological skills in reading Chinese	100
4.3.2.3 Applying Orthographic-Phonology Correspondence (OPC) rules in reading Chinese	101
4.3.3 Semantic aspects of written Chinese	103
4.3.3.1 Understanding that print conveys meaning	103
4.3.4 Conclusion	105

4.4 Literacy development at the formal reading stage	106
4.4.1 Orthographic aspects of written Chinese	106
4.4.1.1 Developing global visual skills in reading Chinese	106
4.4.2 Phonological aspects of written Chinese	112
4.4.2.1 Developing global phonological skills in reading Chinese	112
4.4.2.2 Applying Orthographic-Phonology Correspondence (OPC) rules in reading Chinese	115
4.4.3 Semantic aspects of written Chinese	118
4.4.3.1 Semantic understanding of words	118
4.4.3.2 Detecting semantic radicals	118
4.5 Conclusion	119
4.5.1 Orthographic aspects of written Chinese	119
4.5.2 Phonological aspects of written Chinese	120
4.5.3 Semantic aspects of written Chinese	122
 CHAPTER 5	
 METHOD OF STUDY	
5.1 Rationale of the study	123
5.2 Setting up the experiments	126
5.2.1 Procedure	128
5.3 The Subjects	129
5.3.1 Sample for the pilot Study	129
5.3.2 Sample for the main study	129
5.3.3 Educational background of the subjects	130
 CHAPTER 6	
 STUDY 1: VISUAL DISCRIMINATION OF CHINESE WORDS	
6.1 The Pilot Study	132
6.1.1 What are the units of visual discrimination in Chinese?	132
6.1.2 Subjects	133
6.1.3 Design	133
6.1.4 Materials	134
6.1.5 Procedure	135
6.1.6 Results	136
6.1.6.1 Stroke-complexity	136
6.1.6.2 Unit-complexity	137
6.1.7 Conclusion	138

6.2 The Main Study	140
6.2.1 Difficulty of visual discrimination: number of strokes versus number of stroke-patterns	140
6.2.2 Subjects	140
6.2.3 Design	140
6.2.4 Materials	142
6.2.5 Procedure	142
6.2.6 Results	143
6.2.6.1 Stroke-complexity in the same pairs	143
6.2.6.2 Unit-complexity in the different pairs	144
6.2.7 Conclusion	146

CHAPTER 7

STUDY 2: ORTHOGRAPHIC ACCEPTABILITY OF PSEUDOWORDS AND NONWORDS

7.1 The Pilot Study	148
7.1.1 The formal constraints of written Chinese - the positioning of stroke-patterns	148
7.1.2 Subjects	150
7.1.3 Design	150
7.1.4 Materials	151
7.1.5 Procedure	151
7.1.6 Results	151
7.1.7 Discussion	152
7.2 The Main Study	153
7.2.1 The formal constraints of written Chinese - the graphic characteristics of the stroke-patterns	153
7.2.2 Subjects	155
7.2.3 Design	155
7.2.4 Materials	156
7.2.5 Procedure	156
7.2.6 Results	156
7.2.7 Discussion	159

CHAPTER 8

STUDY 3: CREATIVE SPELLING IN CHINESE

8.1 The Pilot Study	162
8.1.1 Representational functions of semantic radicals and phonological components	162
8.1.2 Subjects	164
8.1.3 The task	164
8.1.4 Materials	165

8.1.5 Procedure	167
8.1.6 Results	168
8.1.6.1 Understanding formal constraints: the position of semantic radicals and phonological components	168
8.1.6.2 Understanding functional constraints: the use of semantic radicals to represent meaning	169
8.1.6.3 Understanding functional constraints: the use of phonological component in reading	171
8.1.7 Discussion	175
8.1.7.1 Understanding formal constraints: the position of semantic radicals and phonological components	175
8.1.7.2 Understanding functional constraints: the use of semantic radicals to represent meaning	175
8.1.7.3 Understanding functional constraints: the use of phonological components in reading	176
8.2 The Main Study	177
8.2.1 Task 1	177
8.2.1.1 Rationale	177
8.2.1.2 Subjects	178
8.2.1.3 Design	178
8.2.1.4 Materials	178
8.2.1.5 Procedure	180
8.2.1.6 Results	180
8.2.1.6.1 Understanding formal constraints: the position of semantic radicals and phonological components	180
8.2.1.6.2 Understanding functional constraints: the use of semantic radicals to represent meaning	182
8.2.1.6.3 Understanding functional constraints: the use of the phonological components in reading	183
8.2.1.7 Discussion	185
8.2.2 Task 2	187
8.2.2.1 The relationship between the positioning and functioning of stroke-patterns	187
8.2.2.2 Subjects	188
8.2.2.3 Design	189
8.2.2.4 Materials	190
8.2.2.5 Procedure	192
8.2.2.6 Results	192
8.2.2.6.1 Understanding the formal and functional constraints: the use of semantic radicals to represent meaning	192
8.2.2.6.2 Understanding functional constraints: the use of the phonological component in reading	195
8.2.2.7 Conclusion	197

CHAPTER 9

STUDY 4 READING CHINESE PSEUDOWORDS

9.1 The Pilot Study	200
9.1.1 The phonological and semantic constraints in reading	200
9.1.2 Subjects	203
9.1.3 Design	203
9.1.4 Materials	204
9.1.5 Procedure	205
9.1.6 Results	206
9.1.6.1 The phonological match	206
9.1.6.2 The semantic match	207
9.1.7 Discussion	208
9.2 The Main Study	209
9.2.1 Task 1 - Pronouncing Chinese pseudowords	209
9.2.1.1 Subjects	209
9.2.1.2 Design	210
9.2.1.3 Materials	212
9.2.1.4 Procedure	212
9.2.1.5 Results	213
9.2.1.5.1 Overall analysis of the use of phonological components	213
9.2.1.5.2 Consistency within child	216
9.2.1.6 Discussion	217
9.2.2 Task 2 - The semantic constraints in choosing meaningful words	219
9.2.2.1 Subjects	219
9.2.2.2 Design	220
9.2.2.3 Materials	222
9.2.2.4 Procedure	222
9.2.2.5 Results	222
9.2.2.6 Discussion	225

10. CHAPTER 10

CONCLUSION AND DISCUSSION 227

10.1 Process of learning to read and write Chinese - holistic versus analytical	229
10.1.1 Unit of visual analysis of Chinese words: the stroke-patterns	230
10.1.2 The formal constraints of written Chinese - the positioning and graphic characteristics of the stroke-patterns	230
10.1.3 The functional characteristics of the stroke-patterns	232
10.1.3.1 The semantic radicals	232
10.1.3.2 The phonological components	233

10.2 The outline of a developmental theory of children learning to read and write Chinese	235
10.2.1 The global identification of the graphic and phonological units	236
10.2.2 Intra-character analysis	237
10.2.3 Predicting the development of reading ability in Chinese	240
10.2.4 Proposed developmental stages in understanding written Chinese at an intra-character level	242
10.3 A theory of instruction	247
10.3.1 What is taught and what is not: implicit and explicit knowledge in learning written Chinese	247
10.3.2 Implications for teaching	249
APPENDICES	251
REFERENCES	259

LIST OF TABLES

TABLE 5-1	THE ADMINISTRATION TIME FOR THE TASKS CONDUCTED IN THE PILOT AND MAIN STUDY	128
TABLE 5-2	NUMBER OF CHILDREN IN EACH AGE GROUP IN THE PILOT STUDY	129
TABLE 5-3	AGE RANGE, SEX DISTRIBUTION AND NUMBER OF CHILDREN IN EACH AGE GROUP IN THE MAIN STUDY	130
TABLE 6-1	MEAN SCORES OF CORRECT JUDGEMENT IN DIFFERENT STROKE-COMPLEXITY BY AGE GROUP	136
TABLE 6-2	MEAN SCORES OF CORRECT JUDGEMENT IN DIFFERENT UNIT-COMPLEXITY BY AGE GROUP	137
TABLE 6-3	MEAN AND STANDARD DEVIATION OF THE NUMBER OF CORRECT JUDGMENTS AT DIFFERENT LEVELS OF STROKE-COMPLEXITY BY AGE GROUP	143
TABLE 6-4	MEAN AND STANDARD DEVIATION OF CORRECT JUDGEMENT IN DIFFERENT UNIT-COMPLEXITY BY AGE GROUP	144
TABLE 7-1	MEAN OF POSITIVE RESPONSES TO PSEUDOWORDS AND 'NONWORDS WITH ILLEGAL POSITIONS' BY AGE GROUP	158
TABLE 7-2	MEAN OF POSITIVE RESPONSES TO PSEUDOWORDS AND 'NONWORDS WITH ILLEGAL GRAPHICS' BY AGE GROUP	159
TABLE 8-1	PERCENTAGE OF PSEUDOWORDS AND NONWORDS IN CREATIVE SPELLING BY AGE GROUP	168
TABLE 8-2	MEAN OUT OF SIX OF PSEUDOWORDS BY AGE GROUP	169
TABLE 8-3	FREQUENCY OF SELECTION OF SEMANTIC RADICALS BY PICTURE	170
TABLE 8-4	MEAN OUT OF SIX AND STANDARD DEVIATION OF CORRECT RADICALS BY AGE GROUP	171
TABLE 8-5	FREQUENCY OF SELECTION OF PHONOLOGICAL COMPONENTS BY PICTURE	172
TABLE 8-6	PERCENTAGE OF PSEUDOWORDS AND NONWORDS IN CREATIVE SPELLING BY AGE GROUP	181
TABLE 8-7	MEAN OUT OF FIVE AND STANDARD DEVIATION OF PSEUDOWORDS	181
TABLE 8-8	FREQUENCY OF SELECTION OF SEMANTIC RADICALS BY PICTURE	182
TABLE 8-9	MEAN OUT OF FIVE AND STANDARD DEVIATION OF CORRECT SELECTION OF SEMANTIC RADICALS BY AGE GROUP	183
TABLE 8-10	FREQUENCY OF SELECTION OF PHONOLOGICAL COMPONENTS BY PICTURE	183
TABLE 8-11	FREQUENCY TABLE FOR SEMANTIC RADICALS MATCHING THE CORRESPONDING PICTURES	193

TABLE 8-12 FREQUENCY TABLE OF CORRECT SEMANTIC RADICALS ALLOCATED TO DIFFERENT POSITIONS	194
TABLE 8-13 MEAN AND STANDARD DEVIATION OF CORRECT POSITIONING OF SEMANTIC RADICALS	194
TABLE 8-14 FREQUENCY TABLE FOR PHONOLOGICAL COMPONENTS MATCHING THE CORRESPONDING PICTURES	195
TABLE 9-1 PERCENTAGE OF CHILDREN'S RESPONSES IN THE PHONOLOGICAL MATCH	206
TABLE 9-2 MEAN SCORES FOR PHONOLOGICAL MATCH BY AGE GROUP	207
TABLE 9-3 PERCENTAGE OF CHILDREN'S RESPONSES IN THE SEMANTIC MATCH	207
TABLE 9-4 MEAN OUT OF TEN FOR CHOOSING CORRECT SEMANTIC RADICALS BY AGE GROUP	208
TABLE 9-5 FREQUENCY OF USING AND NOT USING PHONOLOGICAL COMPONENTS IN PRONOUNCING THE PSEUDOWORDS BY AGE GROUP	214
TABLE 9-6 MEAN OUT OF FIVE AND STANDARD DEVIATION OF CHILDREN USING PHONOLOGICAL COMPONENTS FOR PRONUNCIATION	215
TABLE 9-7 FREQUENCY OF PRONOUNCING PSEUDOWORDS BY DERIVATION AND BY ANALOGY BY AGE GROUP	216
TABLE 9-8 NUMBER OF CHILDREN USING THE SAME TYPE OF STRATEGY FOR PRONUNCIATION IN FOUR OUT OF FIVE ITEMS	216
TABLE 9-9 FREQUENCY OF CHILDREN'S RESPONSES TO THE FIVE PICTURES	222
TABLE 9-10 NUMBER OF CHILDREN CHOOSING THE SAME WORD TYPE IN READING IN FOUR OUT OF FIVE ITEMS	224

LIST OF FIGURES

FIGURE 2.1	ORTHOGRAPHIC STRUCTURE OF WRITTEN CHINESE	28
FIGURE 2.2	COMBINATION OF STROKE-PATTERNS IN FORMING A CHARACTER	31
FIGURE 2.3	SEMANTIC RADICALS AND PHONOLOGICAL COMPONENTS IN IDEOPHONETIC COMPOUND WORDS	35
FIGURE 2.4	EXAMPLES OF THE TWO PART-TO-WHOLE ORTHOGRAPHY-PHONOLOGY CORRESPONDENCE RULES FOR READING CHINESE IDEOPHONETIC COMPOUND WORDS	37
FIGURE 2.5	MEASUREMENT OF CONGRUENCY AND CONSISTENCY IN COMPOUND WORDS	41
FIGURE 3.1	PROBE WORDS AND TEST ITEMS FOR HUANG AND WANG'S (1992) STUDY	62
FIGURE 3.2	SAMPLE MATERIAL ADOPTED IN LAI AND HUANG'S (1988) STUDY	63
FIGURE 4.1	EXAMPLES OF PRESCHOOL CHILDREN'S WRITING	95
FIGURE 6.1	DISTRIBUTION OF THE 20 PAIRS OF PSEUDOWORDS BY UNIT-COMPLEXITY AND STROKE-COMPLEXITY IN STUDY 1	135
FIGURE 6.2	DISTRIBUTION OF THE 'SAME' PAIRS OF PSEUDOWORDS BY STROKE-AND UNIT-COMPLEXITY	141
FIGURE 6.3	DISTRIBUTION OF THE 'DIFFERENT' PAIRS OF PSEUDOWORDS BY UNIT-COMPLEXITY AND NUMBER OF STROKE DIFFERENCE	141
FIGURE 6.4	PERCENTAGE OF CORRECT RESPONSES FOR 'DIFFERENT' PAIRS WITH 2- AND 3- STROKE-PATTERNS BY AGE GROUP	145
FIGURE 7.1	EXAMPLES OF PSEUDOWORDS AND NONWORDS FOR THE PILOT STUDY	150
FIGURE 8.1	PICTURES AND STROKE-PATTERNS FOR THE PILOT STUDY	166
FIGURE 8.2	PERCENTAGE OF CHILDREN USING DIFFERENT STRATEGIES IN PRONOUNCING THE PSEUDOWORDS BY AGE GROUP	174
FIGURE 8.3	PICTURES AND STROKE-PATTERNS FOR TASK 1 IN THE MAIN STUDY	179
FIGURE 8.4	PERCENTAGE OF CHILDREN USING DIFFERENT STRATEGIES IN PRONOUNCING THE PSEUDOWORDS BY AGE GROUP	184
FIGURE 8.5	PICTURES AND STROKE-PATTERNS FOR TASK 2 IN THE MAIN STUDY	191
FIGURE 8.6	PERCENTAGE OF CHILDREN USING DIFFERENT STRATEGIES IN PRONOUNCING THE PSEUDOWORDS BY AGE GROUP	197
FIGURE 9.1	PROTOTYPE AND THE FOUR CHOICES	204
FIGURE 9.2	THE STREET NAMES AND ITS CHOICES	211
FIGURE 9.3	PERCENTAGE OF STRATEGIES IN PRONOUNCING PSEUDOWORDS BY AGE GROUP	214
FIGURE 9.4	PICTURES AND MAKE-UP NAMES FOR THE OBJECTS	221
FIGURE 9.5	PERCENTAGE OF CHOICE OF WORD TYPES BY AGE GROUP	224

Chapter 1

Introduction

Over the last forty years, learning to read and write an alphabetic script has come to be viewed as a cognitive process rather than a rote learning process. Children develop an awareness of the features of their languages and understand it as a system which can be analyzed and manipulated. For example, alphabetic scripts are believed to be learned on the basis of phonological awareness which enables children to learn to read independently (see Goswami and Bryant, 1990 for a discussion on the relationship between phonological skills and learning to read, and Read, 1971; Treiman, 1993 for a description of children's use of letters to represent phonology in English).

Researchers such as Gibson and Levin, 1975; Ferreiro, 1978, 1984, 1985; Ferreiro and Teberosky, 1982; Tolchinsky-Landsman and Levin, 1985 have generally adopted a developmental approach towards the acquisition of the knowledge of the alphabetic scripts. They believe there are qualitative differences, apart from quantitative differences in the stages of learning to read and write. Cognitive concepts of the written language, such as the relationship between print and picture, or between speech and script develop gradually, and the progress is marked by recognizable stages.

It has been shown in several independent investigations that, before children learn to read and write an alphabetic script, they have already developed some understanding of the formal constraints of their written language (see Goelman Oberg & Smith, 1984 for a collection of studies on preschool children and literacy). This understanding provides a basis for learning to read and is predictive of children's reading progress (Ferreiro and Teberosky, 1982; Nunes Carraher and Rego, 1982). It continues to grow as children

learn more about the written language. Ferreiro and Teberosky (1982), for example, showed that young, pre-school children in Argentina who did not yet know how to read used some definite criteria in order to decide whether a graphic pattern could or not be read: they used a minimum quantity criterion, according to which at least three letters are necessary for reading, with single and pairs of letters not being enough for reading; and a variety criterion, which led them to reject a sequence of three identical letters as a sequence that could be read. Karmiloff-Smith (1992) later showed that this criterion was only applied to reading; children considered numbers as proper numbers even when there was only one digit or when the same digit was repeated three times in a string.

Ferreiro and Teberosky (1982) also showed that preschool children who did not know how to read often used not only formal but also a functional criterion when writing: they recognized that the same sequence of letters cannot serve the function of representing two different words. If the children only knew a small sample of letters, they systematically varied their order when trying to write different words. Nunes Carraher and Rego (1982) observed that in Portuguese, before realizing that letters represent phonemes, children seem to attribute to letters the function of representing syllables. They may rely initially on a different set of phonological units, onset and rhyme, before they attribute to letters the function of representing phonemes.

The proposition that learning to read is a process of understanding the written language system has mainly been examined in alphabetic scripts. The central question in this thesis is whether a similar cognitive process is required to learn a 'logographic' script. If the cognitive understanding of the written language system is demonstrated by alphabetic readers being able to establish a relationship between speech and print by matching phonemes to sound, are Chinese readers aware of the presence of phonological units in a word? In addition, the process of learning to read alphabetic scripts is characterized by developmental stages, which shows that the language is being learned as a system rather than in piecemeal fashion. Do Chinese children show the same approach in learning to read?

Written Chinese is regarded by many researchers interested in reading acquisition as a logographic script, using symbols to represent ideas (Temple et al, 1988; Martin, 1972). The relationship between the symbol and sound is described as opaque. Coltheart, for example, suggested that "When faced with a novel word, English readers can pronounce the word by using phonetic cues but Chinese readers cannot use this skill when reading Chinese words" (1978, p. 152). The view of written Chinese as logographic, and thus something which must be learned by rote has been exaggerated so far as to result in the comparison between its acquisition and that of vocabulary by apes: "... there is a human activity that more closely resembles the apes' acquisition of vocabulary. Chinese is written with a system of characters, each of which represents a different word. In order to read and write Chinese, one must learn thousands of characters. Crucially, Chinese children must spend years and years being trained assiduously in order to master the characters. ... We still see a big-brain effect: a Chinese child can learn a whole lot more characters than an ape can learn signs. But the learning is qualitatively similar" (Jackendoff, 1993, p. 139).

In contrast to the 'logographic' view which regards Chinese characters mainly as a combination of strokes, recent psycholinguistic research (Hue, 1992; Tzeng and Hung, 1980; Hoosain, 1991) has brought to our attention the underlying orthographic features of written Chinese. Although the Chinese script looks very different from an alphabetic script, it also contains units with precise formal and functional characteristics. The formal characteristics involve the type and position of the elements which are used in the script; the functional characteristics relate to the type of information which the elements convey. Over 90% of the characters contain graphic units - the stroke patterns - which have particular representational functions: they can either be semantic radicals or phonological components (Zhu, 1987). Studies on Chinese adult readers have provided evidence that these units are recognized and used in reading (Zhang & Simon, 1985; Hue & Erickson, 1988). Adult readers rely on the semantic radicals to identify the meaning of words and pseudowords (Flores d'Arcais, 1992, 1995; Chen, 1993; Chen & Allport, 1995) and on the phonological components for pronunciation (Seidenberg, 1985; Zhu, 1987; Hue, 1992). Furthermore, findings from child studies have also

suggested that learners of written Chinese have to establish different levels of phonological awareness in order to be able to read. Children initially acquire the global phonological awareness by attending to the one-to-one correspondence between a character and a syllable (Lee, 1989; Chan, 1990) and later on pay attention to the phonological component which provides phonological information at an intra-character level (Ho, 1994).

The studies to be reported here were designed to investigate the nature of children's learning of written Chinese. They are based on a view of how written Chinese is learned that radically confronts the idea that children simply learn characters one by one without the support of an underlying system. My hypothesis was that Chinese children, much like children learning an alphabetic script, do not learn written words by rote. They acquire the visual skills needed in word recognition, and they develop an implicit understanding of the formal and functional characteristics of written Chinese. This knowledge is used generatively and allows a learner to recognize, write, and read Chinese pseudowords - that is, graphic patterns which respect the formal and functional constraints of written Chinese but are not real words, and therefore could not have been memorized. If clear evidence to support this hypothesis can be found, it will be possible to conclude that the view that written Chinese is learned by rote cannot be correct.

The literature to be reviewed is discussed in the three chapters that follow. In Chapter two, a brief introduction to the Chinese language is presented. The linguistic structure and some linguistic terms, such as semantic radicals and phonological components, which are often used to introduce the Chinese language but always seem confusing, is explained in detail.

In Chapter three, along with a psychological description of reading processes in adults, I argue that written Chinese is not a logographic language. There is a linguistic unit -

the stroke-pattern - which functions orthographically and linguistically as a letter does in alphabetic scripts. The process of reading Chinese is not holistic, taking the whole character as a logograph; but analytical, referring to linguistic units within a word for semantic and phonological information. The case for a holistic approach is poorly argued in the literature and we could find no evidence to suggest that holistic perception might operate so as to prevent the identification of linguistic components.

In Chapter four, studies that have been carried out investigating the process of children learning to read Chinese is reviewed. It has been assumed that the only way to learn to read Chinese is by rote learning, memorizing characters one by one; after accumulating the knowledge of a few hundred characters, one can then begin to read Chinese. In recent years, research has been carried out in Hong Kong, China and Taiwan investigating the learning process of written Chinese at both the pre-reading stage and the formal reading stage. Contrary to the general belief, evidence has been found that preschool children play a very active role in reading. These findings serve as a starting point to the hypothesis that implicit rules of Chinese can be developed even though children are only taught to read by the 'look-and-say' method.

Chapters five to nine report four experimental studies which were carried out in Hong Kong. Altogether, 236 children participated in the series of studies. The setting of the studies, namely, a game involving a boy on a new planet, involved creating and decoding novel words. The studies were set up to investigate whether the stroke-pattern is the basic unit of visual discrimination in children's reading, and to examine children's awareness of the formal and functional constraints of written Chinese, in relation to the stroke-pattern. A crucial piece of evidence for understanding the written language system is demonstrated by children being able to read pseudowords which have no word-specific knowledge. If Chinese were learnt purely by rote, there would be no way for children to read pseudowords because those words are not real words and they have not been told either the meaning or pronunciation of the words. Furthermore, there would be no way for them to reject nonwords as orthographically

unacceptable because, in theory, any symbol could be a Chinese character if each were learned by rote. If evidence is found that children can read pseudowords, and differentiate pseudowords from nonwords, then it is unlikely that Chinese is learnt by rote. Through these experiments, I show that children gradually acquire knowledge of the orthographic rules of written Chinese by means of the legal positioning of the stroke-patterns; the semantic rules, by gaining an understanding of the presence of semantic radicals and their utilization to get to the meaning; and the phonological rules, by gaining an understanding the presence of phonological components and their utilization for pronunciation.

Chapter 10 summarizes the results, and evidence is presented to show that written Chinese is learnt analytically rather than by rote. The findings that children are analytical in reading and writing at an intra-character level contribute to our understanding of the learning process of written Chinese. Implications for methods of instructions are considered, methods which might be more efficient than current approaches to the teaching of Chinese.

Chapter 2

The Chinese Language

2.1 Introduction

To people who are familiar only with alphabetic script, Chinese script is strikingly different. Today Chinese is spoken by more people than any other languages, and the earliest Chinese writing in existence dates back 3,500 years. The Chinese language has already been discussed in English in the some studies (e.g. Martin, 1972; Wang, 1973; Henderson, 1982; Tzeng and Hung, 1980,1981; Hoosain, 1991; Hu and Huang, 1990). This chapter, accordingly, concentrates on defining terms which may seem confusing to non-Chinese speakers. The chapter begins with some background information on spoken Chinese. Section 3 focuses on the description of written Chinese. There are two kinds of Chinese scripts still in use today. The sophisticated version is adopted in Hong Kong and Taiwan, and the simplified version is practised in mainland China and Singapore. The sophisticated script is adopted in this study. The etymological origin of Chinese word formation is traced in order to clarify the meaning of "logographs", "ideographs", "ideophonetics", etc., terms which have often been used to describe the Chinese language. The orthographic structure of a Chinese word is also clearly defined. The definition of a "semantic radical" and a "phonological component" adopted for this study are also presented in this chapter.

2.2 Spoken Chinese

The phonological structure of Chinese speech units is relatively simple. The syllable is the most important linguistic unit in Chinese because it is the basis writing and many words in Chinese are monosyllables. The syllable can be a single vowel (V) or a consonant plus vowel (CV). There is a strict one-to-one correspondence between a

syllable and a character. Thus a one-character-word consists of one syllable, a two-character-word consists of two syllables; and a three-character-word consists of three syllables. Consonants and vowels are basic phonetic segments combined to form syllables, whereas tones are supra-segmental features that change the pitch levels of syllables. A change in tone brings about a change in meaning and the way the characters are written. For example, [ma] means *to scold* 〈罵⁶〉, [ma] means *hemp* 〈麻⁴〉, [ma] means *mother* 〈媽¹〉 and [ma] means *horse* 〈馬⁵〉¹. No equivalent tonal structure is found in English². Another distinctive feature of spoken Chinese is the massive amount of homophones (i.e. words which differ in meaning but sound the same). In the XIN HUA dictionary, 52 distinct words can be found under the syllable [jiao], including all four different tones. Most individual words usually have no more than 10 homophones including all four tones, and fewer than five, excluding words pronounced with different tones.

There are eight major dialects in spoken Chinese. The vocabulary items used in different dialects for expressing the same concept or object can be totally different. Within one dialect, there may also be different accents. Thus, these dialects can be mutually unintelligible. On the other hand, written Chinese is very consistent across the dialects, the rules in reading have in a way confined the variation in the spoken language. Cantonese is a dialect commonly used in the southern part of China, including Hong Kong, whereas Putonghua is the standard speech adopted in mainland China to unify the language.

¹ The tones are indicated by numbers in Cantonese and by diacritical marks over the vowels in Putonghua. The number of tones in the Chinese language varies from one dialect to another. For example, there are four tones in Putonghua and nine tones in Cantonese.

² In English, variation in tone is used to convey different moods, the meaning of the word being spoken does not change.

2.3 Written Chinese

Writing systems have been qualified as logographic, syllabic or alphabetic, according to the morphemic, syllabic or phonemic representational level of speech. Chinese has long been classified as a logographic language which refers to the use of symbols to represent words or morphemes. Temple (1988) uses the term 'ideographic writing' and describes Chinese writing as a system based on the principle of using a single symbol to represent an idea. The idea is that each symbol stands for a unique whole and it does not give any lexical access via phonological encoding. Tracing the etymological origin of Chinese script provides strong evidence to refute such a belief. Phonology has been taken into account in forming a large proportion of Chinese characters which have a fairly systematic relationship to their sounds.

2.3.1 Six rules for word formation

According to a classical linguistic description of written Chinese, *Liu-shu* 六書, there are six rules of constructing Chinese words: **imitative**, **indicative**, **ideocompound**, **ideophonetic**, **loan** and **transformation**.

The first set of Chinese words were formed by the **imitative** rule which refers to the archaic way of forming written symbols by a schematic 'picture' of real objects. Most of them are single-unit Chinese words formed from individual strokes. For example, the word *fire* 〈火〉 looks like a flame and the two dots represent the sparks coming out from the fire while the word *mouth* 〈口〉 uses a simple square to represent the shape of the mouth. The **indicative** rule refers to the formation of written symbols which can abstractly indicate the meaning. For example, the word *up* 〈上〉 uses a long horizontal line to represent the base, and then uses a vertical line and a small horizontal stroke pointing upwards to indicate the position. The

word *half* 〈 半 〉 uses a vertical stroke to divide two dots and two lines symmetrically. The imitative and indicative rules are in practice difficult to distinguish from each other. Both of them create logographs and ideographs using symbols to represent ideas, and they represent less than ten percent of the Chinese characters still in use today. Instead, they are extensively used as stroke-patterns in the later-formed compound words.

Most of the compound words are formed by the ideocompound rule and the ideophonetic rule. The **ideocompound** rule refers to the composition of words by aggregating the meaning of two components to produce a new word. The word *foolish* 〈 呆 〉 consists of two components, *mouth* 〈 口 〉 and *wood* 〈 木 〉, both of which represent meaning, and together imply a person cannot talk eloquently because of a "wooden mouth". The combination of meaning is similar to the compound words in English, like *blackboard* and *sunlight*. There is no phonological information in words formed by the ideocompound rule.

Phonology has been taken into consideration in forming ideophonetic compound words. The **ideophonetic** rule refers to the composition of words by using a semantic radical to represent its approximate semantic category or semantic field, and a phonological component to represent its pronunciation. There is no counterpart of the semantic radicals in the orthography of other languages. To give an example of the function of the semantic radicals, the stroke-pattern *mouth* 〈 口 〉 is a semantic radical and within this category, some words relating to activities using the mouth can be found, like *eat* 〈 吃 〉, *ask* 〈 問 〉, etc. The semantic and phonological information are represented separately by two different components. For example, the stroke-pattern 〈 馬 〉 means *horse* and it is pronounced as [ma]. It can be made as a phonological component when it is combined with a semantic radical, and the compound word has the same

pronunciation as its phonological component. For example, when the phonological component for *horse* 〈馬〉 is combined with the semantic radical *woman* 〈女〉, we have 〈女〉 + 〈馬〉 = 〈媽〉 which means *mother* and is pronounced as [ma]. When the *horse* 〈馬〉 is combined with *jade* 〈玉〉, we have 〈玉〉 + 〈馬〉 = 〈瑪〉 which means *agate* and again it is pronounced as [ma]. All these words have the same pronunciation as their phonological components.

Transformation and loans are rules of using words, rather than rules of composition. Both rules have also utilized either the semantic and the phonological information, thus they are sometimes called the semantic transformation and the phonetic loan rules. The use of phonetic **loans** applies to situations where a word exists in speech but has not yet acquired the corresponding written form. A word with the same sound - a homophone - is then loaned as the written format for the existing word in speech. It is followed by some graphic alteration such as adding or cutting off some strokes or stroke-patterns. The semantic **transformation** refers to words with the same lexical radical which are to some extent synonyms. Again some strokes/stroke-patterns have been added or cut out to differentiate the words graphically.

Tracing the etymological rule has demonstrated that phonological information has been utilized in at least one type of word formation - ideophonetic compound words. Nowadays, new words are created mainly by the ideophonetic rule which consists of joining semantic and phonological information. Thus the proportion of ideophonetic compound words has been increasing over the centuries (Zhu 1987).

2.3.2 Orthographic units of written Chinese

The Chinese script looks very different from an alphabetic script because the basic orthographic units of the two scripts are very different. Written English has four principal structural levels (letter, word, phrase, and sentence) whereas Chinese has at least five levels (stroke, stroke-pattern, character, word, and sentence). The Chinese orthographic units are introduced in detail in the following sections.

2.3.2.1 A stroke

A stroke is the smallest writing unit in Chinese. There are many types of strokes and the most common ones are dots, lines, and hooks (Figure 2.1). The number of strokes in a word can vary from one to over twenty (e.g. *one* 〈一〉 has just one stroke and *spirit* 〈靈〉 has 24 strokes).

Similar to alphabetic writing, there is a sequence in putting the strokes together in forming Chinese characters. Putting letters in order is simple since all English words are written in linear, left to right direction. But characters are two-dimensional; the sequence might start anywhere, with no set end-point. A fixed sequence of strokes for all characters can prevent any strokes from being left out. It is also believed that characters can be remembered better if the proper stroke sequence is observed.

On the other hand, the function of strokes cannot be equated to letters. They have no phonological function, and they do not correspond to the sound at all. In addition, the identification of the strokes in a character is not as easy as counting letters in an English word. Segmentation of individual strokes in Chinese characters is not always directly visible, and very often it is a matter of writing conventions. For example, it is very difficult to decide how many strokes there are in the one-

character-word *eternity* 〈永〉 because some strokes are linked together (the correct number of strokes for this word is five).

To sum up, a stroke is an important graphic unit but it is not a linguistic unit. Knowing the correct sequence of strokes helps us to recognize the characters quickly and efficiently. The number of strokes can be counted correctly only when we follow a fixed sequence. This skill is crucial for looking up words in the dictionary because all the words are categorized by the number of strokes.

2.3.2.2 A stroke-pattern

Stroke-patterns are formed by putting individual strokes together. In Figure 2.1, there are four stroke-patterns. 〈人〉 consists of two strokes, pronounced as [jan] and means *human being*. 〈半〉 consists of five strokes, it is pronounced as [bun] and it means *half*. 〈口〉 consists of four strokes, pronounced as [hau] and means *mouth*. When the stroke-patterns are used as simple characters in isolation, they account for less than ten percent of the characters in modern usage. However, they have been used extensively as components in the compound characters. For example, the stroke-pattern *mouth* 〈口〉 appears as a component in more than 500 compound characters. Therefore, stroke-patterns are very familiar graphic patterns to readers of Chinese.

Unlike strokes, most stroke-patterns are pronounceable, and all of them have a meaning of their own. However, some of them have lost their sound and/or meaning due to the language's evolutionary change and are only regarded as graphic symbols in modern Chinese. For example, the stroke-pattern 〈宀〉 has a meaning related to roof or shelter but the pronunciation does not exist in modern Chinese any more. It is still in use today, acting as a semantic radical in compound characters.

There is no statistical data about the total number of stroke-patterns, but the commonly use one are estimated to be in the hundreds. The graphic designs of the stroke-patterns have been standardized throughout the years. For example, there are no circular stroke-patterns in Chinese like this 〈^o〉. Dots on top of a character must be either one or two and have to be written as 〈丶〉, 〈丶〉. Dots on the left hand side must be two or three and have to be written as 〈丶〉, 〈丶〉, and four dots are needed horizontally if used at the bottom of a word 〈丶丶〉. In addition, multiple curves have be written as 〈<<<〉 and not as 〈>>>〉. Once the basic stroke-patterns can be identified, it is less difficult to recognize Chinese characters.

Many stroke-patterns occupy fixed positions within a compound character. For example, the three dots on the left (for example 泳) cannot be put on the right side of a word, otherwise a nonword would be formed (for example 泳). There are about twenty stroke-patterns in which the format differs when they are used as simple characters and as semantic radicals in compound characters. For example, the word *water* 〈水〉 is written as 〈氵〉 when it is used as a semantic radical , and the word *human being* 〈人〉 is written as 〈亻〉 when it is used as a semantic radical. In this way, the semantic role of these stroke-patterns can be easily identified.

Figure 2 - 1 Orthographic structure of written Chinese

Orthographic structure of written Chinese	Examples
Strokes ↓ combination of strokes	1. dots 〈丶〉 〈……〉 2. vertical, horizontal and slanting strokes 〈十〉 〈丿〉 〈八〉 3. strokes with hooks 〈戈〉 〈也〉
Stroke-patterns ↓ combination of stroke-patterns	1. human being 〈人〉 [jan ⁴] 2. half 〈半〉 [bun ³] 3. roof/shelter 〈宀〉 - 4. mouth 〈口〉 [hau ²]
Characters* ↓ combination of characters	1. mouth 〈口〉 [hau ²] 2. companion 〈伴〉 [bun ⁶] 3. finish 〈完〉 [jyn ⁴]
Words* ↓ combination of words	1. finish 〈完〉 [jyn ⁴] 2. companion 〈伴侶〉 [bun ⁶ -loey ⁵]
sentences	1. This is my companion. 〔這是我的伴侶。〕

* A single stroke-pattern can be a simple character by itself, e.g. human being 〈人〉, half 〈半〉, mouth 〈口〉

* A single character can be a word by itself, e.g. human being 〈人〉, half 〈半〉, mouth 〈口〉, finish 〈完〉, companion 〈伴〉

2.3.2.3 A character

From the etymological point of view, in Chinese, single characters are classified into two categories, Wen (文) and Zi (字). These two distinctive categories of written characters are constructed by different constituent structures (Hu & Huang, 1990). About 3,500 years ago, the first generation of Chinese characters was formed by putting individual strokes together. The simple strokes were used to form pictographs or ideographs, these now accounting for less than 10% of Chinese characters in modern usage. Wen (文), apart from being used as characters in isolation, are also used as components of the later-formed, compound characters, Zi, which consist of more than one stroke-pattern (字).

Accordingly, the orthographic structure of Chinese characters can be divided into simple forms and compound forms too. A simple character consists of only one stroke-pattern and a compound character is formed from two or more stroke-patterns. A simple character is like the word *mouth* (口) where further decomposition will result in strokes.

Over 90 percent of the characters in modern usage are compound characters and most of the compound characters consist of two stroke-patterns. In Figure 2.1, some examples of compound characters with two stroke-patterns are presented. The stroke-patterns *human being* (亻) and *half* (半) are grouped together to form a compound character *companion* (亻 + 半 = 伴). It is a left-right combination and the stroke-pattern *human being* (人) (亻) has a different format when it is used as a stroke-pattern instead of a character on its own. In another character, the stroke-patterns *roof* (宀) and *beginning* (元) are combined in a top-bottom pattern to form the word *finish* (宀 + 元 = 完). The top-bottom pattern and the left-right pattern are the two most common ways of

combining stroke-patterns in compound characters. Character reading is performed through a multi-directional or 2-D scanning path rather than a one dimensional scanning path.

There are also complicated characters with more than two stroke-patterns. Figure 2.2 demonstrates how the multiple stroke-patterns are combined to form compound characters. The word *pick* 〈采〉 is made by combining two stroke-patterns; the word *once* 〈番〉 combines three stroke-patterns, the word *surname* 〈潘〉 has four stroke-patterns and the word *fence* 〈藩〉 consists of five stroke-patterns. These words seem very complicated, but they are not difficult for Chinese readers to recognise. As shown in Figure 2.2, the complicated characters are formed by simple stroke-patterns which are very familiar to fluent readers. The stroke-patterns always occupy the same positions in the characters, and they are organized along a recursive structure (Huang and Wang, 1992) which means that the stroke-patterns in the brackets are recognized as a Gestalt, rather than as separate units. For example, the two stroke-patterns in *pick* 〈采〉 are regarded as a graphic pattern in the word *once* 〈番〉 and in turn, *once* 〈番〉 is regarded as unique stroke-pattern in *surname* 〈潘〉. It is similar to the graphic combination of words like (ace), p(ace), and s[p(ace)] in English where *ace* and *pace* are taken as a whole unit instead of separate letters.

Every Chinese character is confined in a fixed region, irrespective of the number of strokes in a character. Simple and compound characters all occupy the same prescribed area in regular printed materials. The same does not hold for alphabetic scripts, where the greater the number of letters, the longer the string. Traditionally, characters are written in vertical columns from top down, and columns go from right to left. In modern usage, however, characters are most often written or printed horizontally, line by line, from left to right.

Figure 2.2 Combination of stroke-patterns in forming a character

Number of stroke-patterns	Stroke-patterns	Compound characters
2	𠂇 + 木	采
3	(𠂇 + 木) + 田	番
4	[(𠂇 + 木) + 田] + 彡	潘
5	{[(𠂇 + 木) + 田] + 彡 } + 艹	藩

2.3.2.4 A word

In English, a word is identified in terms of space at either end of a letter string in a text. In Chinese, a word can be written by one character, two characters, or three characters with no graphic linkage between the characters to indicate the boundary of a word. In the classical literary style of writing, a Chinese character normally represents a word (e.g. 伴 meaning *companion*). However, in the modern colloquial style of writing, many words are written with two characters (e.g. 伴侶 meaning *companion*) although the most frequently used ones tend to be of one character only. Of the 60 words ranked highest in usage (Suen 1986, cited in Hoosain 1991), only two are two-character words and the rest are one-character words. A single character is sometimes described as a single "morpheme" in a multi-character-word, but it is a word by itself when it is used independently (Figure 2.1). The term 'a character' has been widely used in Chinese psycholinguistic studies and it is, in fact, a 'one-character-word'. In this study, 'a single-character-word' is referred to as 'a word' in brief. Multi-character-words are specified by 'two-character-word' and 'three-character-word'.

As for the size of the Chinese lexicon, about 3,000 characters account for 99% of the total frequency of Chinese words used in modern books and newspapers in

mainland China, according to the Foreign Languages Press Beijing (1989). According to another survey in Taiwan, there are about 4,500 characters and about 40,000 words in daily use (Liu, Chuang and Wang, 1975).

There are four ways of classifying single-character words in Chinese dictionaries, and the semantic radical index is the most popular sorting system. Chinese words are listed according to semantic radicals and semantic radicals are arranged according to the number of strokes.

2.3.2.5 A sentence

A sentence consists of words. Characters in text are evenly spaced with no extra space between words. Readers have to segment character strings into words for comprehension. In Figure 2.1, there are five words in the sentence "This is my companion" - *This* 〈這〉 *is* 〈是〉 *my* 〈我〉 〈的〉 (the second word changes the first word *I* 〈我〉 to possessive *my*) *companion* 〈伴侶〉. Segmenting larger segments in Chinese sentences is the same as the English text. Clause and phrase boundaries are indicated by commas and semicolons, sentence boundaries are marked by full stops.

2.3.3 Orthographic constituents of a ideophonetic compound word

About 90% of Chinese characters are compounds (Zhu, 1987), in which the character is composed of stroke-patterns. A ideophonetic compound character consists of two linguistic components - the semantic radical which indicates the meaning and the phonological component which indicates the sound. In contrast to English orthography, which specifies both semantic and phonological information across the whole letter array, in Chinese, semantic and phonological information are encoded in separate orthographic components. For example, in the word

companion 〈伴〉 [bun³], the stroke-pattern *human being* 〈亻〉 is the semantic radical which gives a clue to the meaning of the word or its related category (the *companion* is a person), and the stroke-pattern *half* 〈半〉 [bun] is the phonological component which gives a clue to the pronunciation of the word.

2.3.3.1 The semantic radical

The semantic radical refers to any of the 214 indices listed in the Chinese dictionary when these components occur in the radical position in a character. Many semantic radicals are logographs or ideographs by themselves, and they are used as the semantic components in the characters. Some of them are simple characters in their own right, and some are unpronounceable graphic patterns. In either condition, semantic radicals are never pronounced in compound words.

There is no counterpart to the semantic radical in the orthography of other languages. Semantic radicals originated from the idea that words were to express meaning. In the logographic stage of written language development, Chinese characters were represented by pictorial symbols. The idea of expressing meaning then developed into another stage, using semantic radicals for categorization. For example, the categories for 'human activities' consisted of activities using eyes, ears, feet, hands, and the whole body. In a word like *eat*, we expect the semantic radical *mouth* to be there, so the word *eat* would be written like this "mouth + xxx" (where xxx can be a phonological component without any meaning) and *kick* would be written like this "foot + xxx". There are other semantic categories which are related to objects such as nature (e.g. earth, wind, etc.), flora (e.g. grain, rice, etc.), fauna (e.g. horse, pig, etc.), etc. In the same way, we expect a kind of flower, for example *chrysanthemum* to be written as "flora + xxx". In this way, the semantic radical gives a semantic clue to the meaning of the characters.

Sometimes the semantic implication of the semantic radical is transparent, and sometimes it is opaque. For example, the semantic radical *female* 〈女〉 in the word *mother* 〈女〉 + 〈馬〉 = 〈媽〉 is straightforward because it has clearly indicated the gender category. However, the semantic relationship becomes unclear in the word *good* 〈女〉 + 〈子〉 = 〈好〉 unless the Chinese culture is taken into account. In Chinese culture, it is often believed that for a woman to have a son is *good*, and that is why the word is made up of *female* 〈女〉 and *son* 〈子〉. Semantic radicals usually have a more direct and relevant relationship with verbs and nouns and a less straightforward relationship with adjectives, connectives and function words.

The semantic radical always occupies a standard position in a character. For the left-right pattern compounds, the semantic radicals are usually on the left. For example, the semantic radical *human being* 〈亻〉 is at the left hand side of the word *companion* 〈伴〉. For the top-bottom patterns, the semantic radicals are usually on top. For example the semantic radical *roof* 〈宀〉 is on top of the word *finish* 〈完〉. Relatively fewer semantic radicals are found at the right hand side of left-right patterned characters or at the bottom of top-bottomed patterned characters.

In some cases, a component may be a semantic radical in one character and a phonological component in another. However, this is easily distinguished because the same component used as a semantic or as a phonological component always occurs in different positions depending on its semantic radical or phonological status. For example, the stroke-pattern *horse* 〈馬〉 [ma⁴] is listed as a semantic radical in the dictionary. It appears as a semantic radical in the compound word *donkey* 〈馬〉 + 〈盧〉 = 〈驢〉 [lou⁴] and the stroke-pattern 〈馬〉 is on the left hand side which is its legal radical position. The same stroke-pattern *horse* 〈馬〉

[ma] can also appear in a non-radical position, such as in the word *mother* 〈女〉 + 〈馬〉 = 〈媽〉 [ma⁵], and 〈馬〉 is on the right hand side of the word which is its legal phonological position (Figure 2.3). Therefore, even though some stroke-patterns can appear in different positions, their linguistic status is decided by their positions within the character. In addition, there are plenty of semantic radicals with only one fixed graphic position. For example, the semantic radical 〈氵〉 has to be placed on the right side of a word 〈泳〉, otherwise it yields an illegal word (i.e. a nonword), e.g. 〈氷〉.

Figure 2.3 Semantic radicals and phonological components in ideophonic compound words

Semantic radicals	Phonological components	Ideophonic compound words
馬 [ma ⁵]	盧 [lou ⁴]	馬盧 [lou ⁴]
女 [noey ⁵]	馬 [ma ⁵]	媽馬 [ma ¹]

The semantic radical is never pronounced and the phonological information can only be found in the non-radical component. Using the same example, when the stroke-pattern *horse* 〈馬〉 [ma⁵] is used as a semantic radical, as in the word *donkey* 〈馬〉 + 〈盧〉 = 〈驢〉 [lou⁴], the stroke-pattern 〈馬〉 is not pronounced. The phonological component is in the stroke-pattern 〈盧〉 [lou⁴] on the right hand side of the word. When the same stroke-pattern *horse* 〈馬〉 [ma⁵] appears as a non-radical component, as in the word *mother* 〈女〉 + 〈馬〉 = 〈媽〉 [ma¹], it becomes the phonological component and provides the phonological clue to read the word (Figure 2.3).

Apart from the legal positioning of the semantic radicals, some stroke-patterns acquire a different format when they are used as semantic radicals in order to facilitate the easy identification of compound words. For example, the stroke-pattern *human being* 〈人〉 is written like this 〈亻〉, and *water* 〈水〉 is written as 〈氵〉.

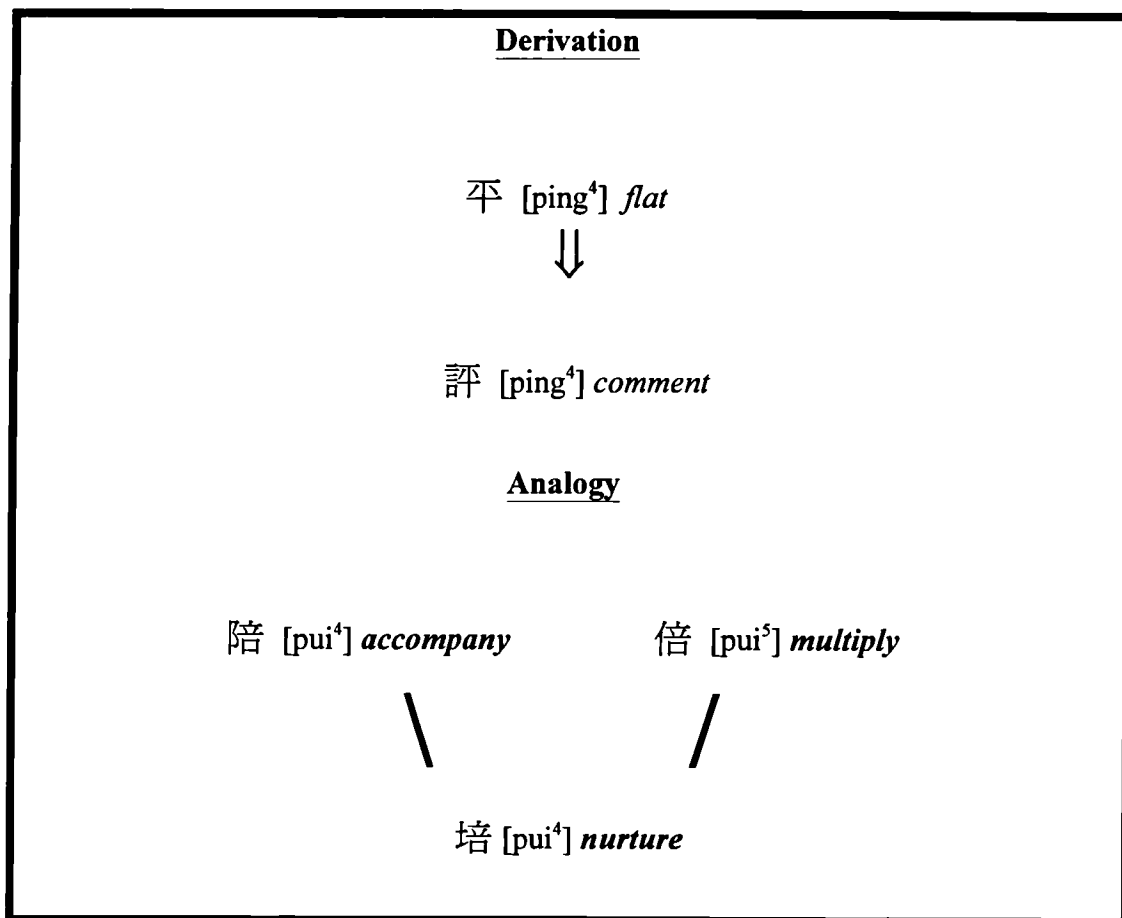
2.3.3.2 The phonological component

There are about 800 phonological components which give clues to the sound of ideophonic compound words. It is not always clear, even for skilled readers, which is the phonological component in a word. However, Chen (1993) has proposed that, as a rule of thumb, the non-radical component is often the phonological component. As there are habitual positions for semantic radicals, by default, there are also habitual positions for phonological components. Thus, the distinction between the semantic radicals and the phonological components are structurally defined by their habitual positions in the characters.

Again, phonological components can be simple words in their own right or unpronounceable graphic units. If the phonological component is a real word, pronunciation can be achieved by derivation, that is, the pronunciation of the whole word from the sound of its phonological component (e.g. the sound of the word *comment* 〈評〉 [ping⁴] can be directly obtained from the sound of its phonological component *flat* 〈平〉 [ping⁴]; or by analogy, that is, pronunciation is derived from words sharing the same phonological components. If the phonological component is an unpronounceable stroke-pattern, pronunciation can only be achieved by analogy. For instance, 〈倍〉 is an unpronounceable phonological component, the sound of the word *nurture* 〈培〉 [pui⁴] can be deduced from the words *accompany*

〈陪〉 [pui⁴] and *multiply* 〈倍〉 [pui⁵] by analogy as they share the same phonological component (Figure 2.4).

Figure 2.4 Examples of the two Part-to-whole Orthography-Phonology Correspondence rules for reading Chinese ideophonetic compound words



2.3.3.2.1 The Orthography-Phonology Correspondence Rules

There are at least two possible ways that the pronunciation of a word could be reached via the phonological component. It can be reached either through:

1. Derivation : pronunciation of the whole word directly derived from the pronunciation of its phonological components, where the phonological components are pronounceable stroke-patterns ; or
2. Analogy : pronunciation of the whole word is deduced via analogy with other words sharing the same phonological component, where the phonological components can be pronounceable or unpronounceable.

Chen (1993) called the two ways of pronunciation via the phonological components the "orthography-phonology correspondence" (OPC) rules. These OPC rules are different from the grapheme-phoneme correspondence (GPC) rules in English as the method of phonological notation in reading English is "atomistic", whereas it is "holistic" in reading Chinese (Wang, 1981). Chen used the term "part-to-whole" to explain the relationship. For words in alphabetic languages, each orthographic unit (i.e. a letter) has a unique phonetic value. The pronunciation of a regular word is arrived at by assembling the phonetic values of all the orthographic units in a word. In contrast, the pronunciation of a Chinese word can be derived from the phonological value of just one orthographic component, i.e. the phonological component of the word. It is noteworthy that the pronunciation of the phonological component and the whole word are always single syllables. Thus, the phonological component, as a "part" of the word, encodes, or specifies the syllabic pronunciation of the "whole" word.

2.3.3.2.2 Phonological regularity of the Chinese orthography

The phonological regularity of an orthography is usually defined as the congruency of the pronunciations of the constituents (letters or letter strings in alphabetic languages) of words. So phonological regularity of Chinese is the congruency of the phonological value of the phonological component in compound words.

Regarding the phonological regularity of ideophonic compound words, Zhou (1980) reported that the predictive accuracy of the pronunciation of an ideophonic compound word from its phonological component is about 39%. Similarly, according to Zhu's (1987) statistical analyses, the predictive accuracy of the pronunciation of an ideophonic compound from its phonological component is about 44% based on the same criteria used by Zhou (1980). However, the predictive accuracy drops to 19% if frequency of usage is taken into consideration. In other words, ideophonic compounds are more "phonologically regular" in low frequency words than high frequency ones. Therefore, Zhu concluded that the phonological cueing function of the phonological components in reading Chinese ideophonic compounds is weak.

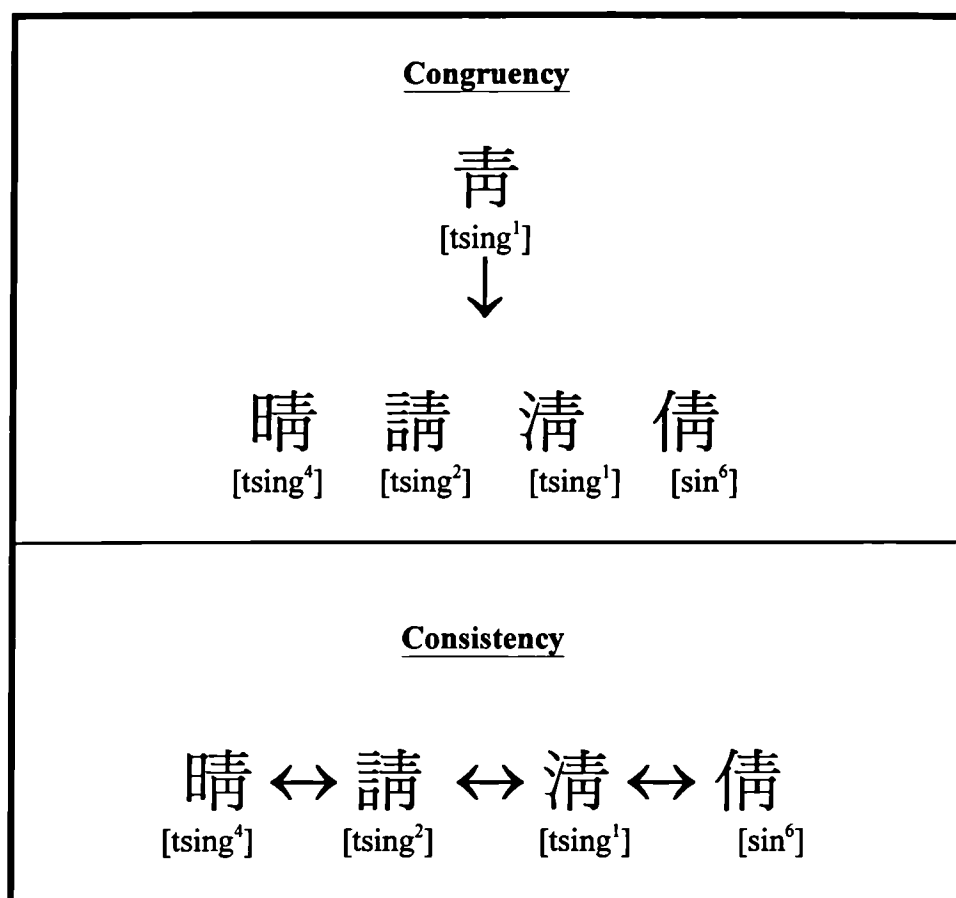
To draw such a conclusion is assuming that "derivation" is the only rule for decoding ideophonic compounds (Figure 2.4). Zhu (1987) and Zhou (1980) have assumed that the pronunciation of ideophonic compounds can only be derived directly from their phonological components. However, compound words can also read by analogy, i.e. by relating to the sound of other words having the same phonological components.

In fact, these linguistic studies have only provided congruency indices, i.e. indices of the degree to which pronunciation of the ideophonic compound is congruent with the pronunciation of its phonological component in isolation. This is an index of intra-word relations. The congruency index is a useful indicator of "phonological

regularity" if readers derive word pronunciation directly from that of its phonological component, i.e. via the derivation rule. However, the congruency index can be calculated only when the phonological component is a real word which can be pronounced in isolation.

Fang, Horng, and Tseng (1986) and also Chen (1993) have proposed another useful, but often neglected, index to measure "phonological regularity" of ideophonic compounds. This is the consistency index. It is an index of the degree to which compound words with the same phonological component are consistent in pronunciation with each other, regardless of possible tonal discrepancy. This is an index of inter-word relations (Figure 2.5). The consistency index in Chinese is important if readers derive pronunciation of a word by making an analogy with other words sharing the same phonological component especially when the phonological component is an unpronounceable graphic unit. Thus the phonological regularity of Chinese words should be evaluated according to at least two indices : the intra-word congruency index and the inter-word consistency index.

Figure 2.5 Measurement of congruency and consistency in compound words



2.3.4 Summary

The following ten points summarize the distinctive orthographic structure of written Chinese:

1. A stroke is the smallest graphic unit.
2. A stroke-pattern is made up of strokes. Stroke-patterns are familiar graphic patterns which are used as building blocks for over 90 % of the characters.
3. A simple character consists of one stroke-pattern, a compound character consists of two or more stroke-patterns. There are some orthographic rules which regulate the form and positioning of the stroke-patterns.

4. A word may consist of one character, two characters, or three characters. There is a one-to-one correspondence between a character and a syllable. Thus a one-character-word consists of one syllable, two-character-word consists of two syllables and so on.
5. A ideophonic compound character consists of a semantic radical and a phonological component. The semantic and phonological information are specified by separate orthographic components.
6. There are altogether 214 semantic radicals in written Chinese. They provide semantic clues to the words.
7. The phonological components refer to the non-radical components in the ideophonic compound characters. Phonological information may be embedded in these components.
8. The orthographic-phonology correspondence rules describe two ways of pronouncing of a Chinese word via the phonological components.
9. Pronunciation of a word can be reached by analogy and derivation via the phonological components.
10. Phonological regularity of Chinese can be measured by consistency indices and congruency indices.

Chapter 3

The Skills of fluent Chinese readers

3.1 Introduction

In this section, a number of studies are to be reviewed in order to answer the following questions: 1) How do adult readers read Chinese? What is the role of orthographic, semantic and phonological information in reading Chinese? 2) Since the Chinese script is different from an alphabetic script, does this difference result in different reading processes across the two writing systems? Are there different representation systems for alphabetic and non-alphabetic systems? Extensive research has been carried out to examine the reading process of alphabetic scripts, and various models, such as dual-route models (Colheart, 1980), lexical analogy models (Glushko, 1979, McClelland & Rumelhart, 1981)) were introduced to explain the complicated phonological processing and the possible factors in influencing the pronunciation known words and novel words.

Research that has been carried out in tracing the reading processes of Chinese is mainly based on the dual route models. Colheart (1978) envisaged two routes in reading an alphabetic language: a direct route, which accesses a logogen type of unit; and a grapheme-to-phoneme conversion route. He also claimed that Chinese script is not suitable for grapheme-phoneme conversion because it is ideographic and no counterparts of the English conversion rules exist in Chinese. Although Chinese cannot be read by graphic-phoneme conversion, the two routes to achieve reading in Chinese exist. Yin and Butterworth (1992) conducted a neuropsychological study by examining eleven brain-damaged patients in mainland China about their ability to read aloud Chinese characters. Evidence is presented that one group of patients (deep dyslexic patients) suffer a selective deficit to a routine that maps whole characters onto pronunciations, and a second group (surface dyslexic patients) suffer an impairment to a sublexical routine that utilises the phonological components, that make up words.

All deep dyslexic patients were completely unable to read pseudowords aloud. On the other hand, the surface dyslexics were able to read some pseudowords by referring to the phonological information provided by the phonological components. The study clearly demonstrated that both routes exist in reading Chinese, and in order to learn to read novel words, the ability to read characters at a sublexical level is of an advantage.

Other studies examined how Chinese words are processed in the form of logographs (Perfetti and Zhang, 1991; Cheng, 1992) or at a sublexical level (Seidenberg, 1985; Zhu, 1987, Hue, 1992). They were termed as "holistic approach" and "analytical approach" in this study. The 'holistic approach' treats characters as logographs, neglecting the presence of the semantic radicals and phonological components. The 'analytical approach' acknowledges the semantic radicals and phonological components as functional linguistic units in 90% of the characters and their role in providing orthographic, phonological and semantic information in word recognition. Before reviewing the literature, I will present a brief introduction to the paradigms used in Chinese reading research.

3.2 Methodological issues

The most common methods devised to examine reading strategies of adult Chinese are: a) lexical decision tasks, b) naming tasks and c) interference tasks.

Lexical decision tasks use real words and pseudowords as stimuli and subjects are asked to decide if the target word/words are words. Chinese pseudowords consist of familiar stroke-patterns allocated to legal positions. They are generated according to the orthographic structure, and theoretically, they could have meaning and be pronounced, it is 'accidental' that they have not been adopted as real words.

The use of pseudowords in lexical decision tasks requires that they be systematically constructed in order to tap the linguistic knowledge of subjects, as has been done in studies in English (Marsh, 1980) and in Chinese (Fang, 1986; Chen, 1993; Ho, 1994). They are acceptable as words but they do not have lexical status. When they are generated in this way, they can be used to examine the strategies of readers in reading and understanding the meaning of unknown words.

On the other hand, in some Chinese studies (Cheng & Shih, 1988; Cheng, 1992), pseudowords are treated as nonsense words in lexical decision tasks. The researchers have overlooked the possible semantic and phonological information provided by the pseudowords which are the unwanted bit of information if pseudowords are only treated as nonsense words. Pseudowords are not nonsense words because they are formed according to the orthographic rules of written Chinese, and they are likely to lead to ambiguity of results if they are treated as nonwords.

In contrast to pseudowords, nonwords are constructed by violating the orthographic structure and they are not regarded as words at all. Nonwords are more similar to unpronounceable letter strings in English or strings containing non-letter forms (digits, reversed letters, etc.).

Interference methods are used in order to examine the various strategies the subjects adopt in reading. Barron and Baron (1977) created phonological interference by asking children to repeat the word 'double' throughout the task, so that they are effectively blocking out all other phonological and articulatory activity. If the task involves a phonological strategy, then the child's performance will be hindered by having to say 'double, double'. But if the task is done purely visually this interference should have little or no effect. Perfetti and Zhang (1991) used a backward masking method to investigate whether Chinese readers use semantic, phonological or graphic information in word recognition by assessing which kind of

masking effect (semantic, phonological and graphic) produced the greatest interference in recalling the correct target words.

Another method which is widely used in Chinese reading research is the naming task (Seidenberg, 1985; Zhu, 1987; Hue, 1992). Subjects are required to name the stimuli. The perceptibility of a stimulus is diminished to a point where errors are ensured by displaying it extremely briefly, and sometimes, by following it immediately with an irrelevant interfering stimulus. The purpose of the method is to detect readers' use of their knowledge of stimulus structure in order to facilitate word recognition.

3.3. Phonological aspects of reading Chinese

3.3.1 Holistic approach

The following studies investigated the reading processes of Chinese using character as the basic phonological unit. Cheng (1992) believed that, despite the fact that no grapheme-to-phoneme rules exist in logographic Chinese words, reading Chinese still requires phonological mediation, and words are processed predominantly by phonological codes rather than by visual codes. The mediation is thought to be based on character-sound correspondence which is well developed through years of extensive practice. He compared the phonological and visual processing of Chinese, predicting that phonological clues are more useful than visual clues in a lexical-decision task. Subjects were to judge on each trial whether or not a target word, preceded by a clue, was a real Chinese word. The variables in the experiment were (1) visual and phonological similarities between the clue and the target, and (2) the asynchrony between the onset of the clue and that of the target.

Twenty-four native speakers of Chinese participated in this experiment. Stimuli consisted of 224 legal Chinese words and 96 pseudowords. The stimuli were used

to form 64 word - word pairs, and 64 word - pseudoword pairs with the first item serving as a clue and the second as a target for pair presentation. There were 4 conditions with 16 pairs in each condition; (1) visually-similar and phonologically-identical pairs (e.g. 培[péi] - 陪 [péi]), (2) visually-similar and phonologically-dissimilar pairs (e.g. 讀[dú] - 續 [xù]), (3) visually-dissimilar and phonologically-identical pairs (e.g. 那 [nà] - 納 [nà]) and (4) visually-dissimilar and phonologically-dissimilar pairs (e.g. 補[bǔ] - 譯 [yì]) which were treated as control pairs. Each word comprised two components - the semantic radical and the phonological component. The pseudowords were formed by replacing the radical component by another radical component in order to form a pseudoword. The position of the components were in their legal positions.

For each trial, the subjects were required to focus attention on the fixation point immediately upon receiving an auditory warning signal. The fixation point was exposed for two seconds and then was replaced by a clue followed by a target. The target appeared at the time the clue was terminated. The subjects were to judge as quickly as possible whether or not the target was a real word by pressing an appropriate response key. The subjects were timed from the onset of the target to the response onset. All the trials were individually administered.

Results showed that word pairs that were phonologically identical in both visually-similar and visually-dissimilar pairs were judged faster than control pairs sharing no similarity of either dimension. However, when the pairs were visually similar only, the decision times were longer than that for the controls. The author concluded that the lexical access was similar to English, and the result was comparable to Meyer et al's (1974) study, as the words were obviously recoded phonologically and acted as a prime to affect the lexical decision to a subsequent target word.

In Meyer et al's (1974) lexical decision task, they found that decision times were influenced by the degree of visual and phonological similarity shared by the strings of words. Specifically, they found that subjects responded more rapidly to word pairs which were both visually and phonologically similar (e.g. BRIBE-TRIBE, FENCE-HENCE) than control pairs sharing no similarity of either dimension (e.g. BRIBE-HENCE, FENCE-TRIBE). However, when the pairs were visually similar only (e.g. FREAK-BREAK, COUCH-TOUCH), decision times were significantly slower than those for appropriate controls (e.g. FREAK-TOUCH, COUCH-BREAK). These results were interpreted by Meyer et al to mean that pair members are processed sequentially and both require phonological recoding prior to lexical access. When pair members are visually similar, the phonological recoding of the first item may influence the manner in which the second item is recoded. That is, there is a bias to use the same GPC rules on the second item as were employed in the recoding of the first. If the two graphemic codes map on to similar phoneme codes, the lexical decision can be facilitated. On the other hand, if the two grapheme codes should map onto different phonemic codes, the bias will lead to an incorrect phonemic representation, which is finally subject to suppression, thereby slowing down lexical decisions.

Doubt arises when Cheng's (1992) results are compared with the English study. In Meyer et al's (1974) study, the response to the phonologically similar pairs was faster because it was facilitated by the graphemic-phoneme rule and in return became a hindrance in the response to the visually-similar and phonologically different pairs. Chinese cannot be read by grapheme-phoneme rules because these elements do not exist in Chinese words. Cheng has not explained the reason behind the facilitation and hindrance of phonologically-identical and phonologically-dissimilar pairs.

A closer look at the orthographic constituent in the word pairs may provide us with some explanation. In the visually-similar pairs, the phonological components (on the right side of the words) were the same and the semantic radicals (on the left side

of the words) were different (1) the visually-similar and phonologically-identical pairs (e.g. 培[péi] - 陪[péi]), (2) visually-similar and phonologically-dissimilar pairs (e.g. 讀[dú] - 續[xù]). The phonologically identical pairs (1) were judged faster with very few mistakes and the phonologically dissimilar pairs (2) were judged slower with a lot of mistakes. The encoding bias may be due to the identical phonological components in the pairs which encouraged the subject to use the orthographic-phonology correspondence (OPC) rules. The subjects encoded the first word phonologically by analogy, referring to the phonological components, and were biased toward using the same rules to recode the second word as it had the same phonological component as the first word. The identical phonological components had facilitated the lexical decision in the visually-similar and phonologically-identical pairs (1). However, it had led to incorrect decisions in the visually-similar and phonologically dissimilar pairs (2) as not all the words with identical phonological components are pronounced the same in Chinese.

Although the orthographic-phonology correspondence rule can explain the phonological processing of the visually similar pairs with identical phonological components, it is still doubtful why the visually-dissimilar and phonologically-identical pairs were processed the fastest in this experiment.

In another study, Perfetti and Zhang (1991) also investigated if the phonological process of Chinese, which is regarded as a logographic language, is similar to English. Perfetti et al (1988) had reported experiments on backward visual masking that were designed to reveal early-occurring prelexical phonemic processes in English word identification. The important result in the experiments was a reduction in the deleterious effect of a pseudoword mask that followed a briefly presented target word when the mask shared phonemic information with the target. In an experimental trial, the subject might see the sequence *rate* - RAIT - ####, *rate* - RALT - ####, or *rate* - BUSK - ####. The first word is a target word presented for



around 30 ms, and a pattern mask ends the sequence and remains on while the subject writes the target word. In the example above for *rate*, there is a phonemic mask, a grapheme mask, and a control mask, respectively. The phoneme mask produced better target identification than did the graphemic mask, which controlled the number of target letters appearing in the mask, and both produced better performance than the control mask.

A replication study on Chinese was carried out to investigate the reading process of Chinese under conditions that had produced evidence for early phonemic activation in English. A semantic mask was added expecting a semantic masking effect.

Forty native speakers of Chinese participated in the experiment. The materials consisted of 34 Chinese target words and their associated masks, four for each target. The masks included (a) a graphic mask, which has substantial visual overlap - because of the presence of an identical component with the target word (b) a phonological mask, which was homophonic to the target word but was not visually similar to it, (c) a semantic mask, which was closely related in meaning to the target word but was neither visually nor phonologically similar to it, and (d) a control mask, which had little graphic, phonological, or semantic similarity to the target. For example for the target word *watch* 視 [shì], the homophonic mask is *matter* 事 [shì], the graphic mask is *now* 現 [xiàn], the semantic mask is *see* 看 [kàn] and the control mask is *clear* 清 [qīng].

Each trial consisted of a briefly exposed target word followed immediately by a mask, that displaced it. The mask was exposed for 30 ms and was displaced by a pattern mask. The pattern mask remained on while the subject wrote down the target word. The results indicated that the only effect was the graphic mask, which produced better target identification than the other three mask types, which did not

differ from each other. The graphic mask effect was significant when the target was exposed for 30 ms, 40 ms, 50 ms, 60 ms and 70 ms. On the other hand, there was no evidence for a phonological mask reduction effect. The homophonic effect failed to reduce the deleterious effect of word masking in Chinese word identification, a result which is exactly the opposite to the data collected with English subjects (Perfetti et al 1988).

All the graphic masks in this study shared either the semantic radicals or the phonological components with the target words, and it is likely that the components had provided some semantic and phonological clues to the target words. This may be the reason for the facilitation of target words. For example, the target word *watch* 視 [shì] and the graphic mask *now* 現 [xiàn] shared the same phonological component (on the right of the words). The additional phonological information may have helped to deduce the deleterious effect of the graphic mask.

One drastic drawback in the studies (Cheng, 1992; Perfetti and Zhang, 1991), treating Chinese characters as logographs, is that they disregard the phonological components within the words. In English, phonological information is encoded across the whole letter array and blending the pronunciation from the letters is in most cases restricted by graphemic-phonemic rules. In Chinese, the phonological information is encoded in the phonological components in over 90 percent of the compound single-character-words. It is impossible to exclude the effect of the phonological component in examining the phonological processing of compound words. The psychological reality of referring to the phonological components for pronunciation had been shown in the lexical-decision task in Cheng's (1992) study, affecting the speed in deciding whether a pair of words were legal or not, and also in the masking task in Perfetti and Zhang's (1991) study, in providing additional phonological information to the graphic masks. Therefore, results obtained from phonological studies which failed to take into account the phonological components cannot provide reliable information due to these confounding factors. A closer look

into the phonological components may provide a better understanding of phonological processing in reading Chinese.

3.3.2 Analytical approach

Seidenberg's (1985) study was one of the earliest attempts to consider phonological components for pronunciation. He investigated whether the phonological information coded in the phonological components of compound Chinese words was utilized for pronunciation; and the extent to which this phonological information influenced reading of both Chinese and English words at different levels of frequency. There were two experiments in this study. In one experiment, the facilitation of the phonological components in reading Chinese was investigated by asking Chinese subjects to read aloud (1) words with the phonological components providing useful phonological information, (2) words with no phonological components (i.e. simple characters) and (3) words with phonological components that did not provide useful phonological information.

Twenty-two Chinese adult readers were asked to read aloud four types of tachistoscopically presented words: high and low frequency phonograms and non-phonograms. By "phonograms" the author meant those compound words containing phonological components that provide information relevant to pronunciation. By "non-phonograms" he meant single form words (in which the word is one holistic stroke pattern with no phonological component) or compound words where the phonological components do not provide reliable information for pronunciation.

The results confirmed the importance of the phonological components for pronunciation. By manipulating the information provided by different phonological components, a significant difference was found in reading low frequency words. Phonograms were named faster than non-phonograms in low frequency words. In the case of high frequency words, the naming latencies for phonograms and non-phonograms did not differ significantly.

In another experiment, the English script was investigated to examine the utilization of phonological information. 48 English adult readers were asked to read aloud four types of words: high and low frequency regular and irregular words. The procedure was the same as in the first experiment. It was found that with low frequency English words, exceptions took longer to name than regular words. Naming latencies for high frequency regular and exception words did not differ significantly.

Seidenberg concluded that the phonological information encoded in some Chinese words is actually used in reading. In both Chinese and English, readers use lexical information to get at the sound of familiar words but use phonological information to read unfamiliar words.

Even though the utilization of phonological information in both English and Chinese has been confirmed in this study, Seidenberg has only acknowledged the facilitating effect of the phonological components. The phonological components may hinder identification when the pronunciation of the phonological component and the word are not congruent. In this study, the simple form words were mixed with the compound words in the non-phonograms. The hindrance effect was not taken into consideration.

Zhu (1987) investigated the congruency effect of phonological components in naming high frequency and low frequency Chinese words. The congruency effect

refers to the sound congruency between the word and its phonological component. If Chinese readers refer to the phonological components for pronunciation, then the congruency effect should affect the speed in naming words. Chinese adult readers were asked to name two types of visually presented words: (1) congruent words (the pronunciation of the word is the same as its phonological components), and (2) incongruent words (the pronunciation of the compound word is totally different from that of its phonological component) or simple words with only one stroke-pattern. In addition, the two types of words varied in different levels of frequency.

It was found that in the two lowest frequency ranges, the naming latencies for congruent words were significantly faster than for words giving no sound clues or incongruent sound clues but that the naming latencies for the two types of words did not differ significantly in the three higher frequency ranges. Zhu concluded that the phonological components of compound words does give some phonological clues for naming unfamiliar Chinese words. However, this study has two limitations: (1) as incongruent ideophonetic compounds and single form words were grouped together as a single stimulus type, it is unclear whether the former, which provide inconsistent phonological clues, result in faster or slower naming latencies than the latter, which provide no phonological clues. (2) The author has assumed that pronunciation can only be achieved via the phonological components by derivation only, which is not true, because compound words can also be pronounced by analogy, i.e., by relating to the sound of other words having the same phonological components. Thus these two factors have to be taken into account in future research.

Hue's (1992) naming latency study also investigated the congruency and consistency effects in reading and he has improved Seidenberg's (1985) study by manipulating the word frequency and partitioned out the simple characters with no phonological components as a unique word type. Ten subjects were involved in the experiment. Six sets of words were selected by crossing frequency (high and low) and word type (congruent, incongruent and unique). Results showed that high frequency words

were read much faster than low frequency words, and among the low frequency words, the congruent words were read the fastest and the incongruent words were read the slowest. The long latency for low frequency irregular word might be due to the competition between the pronunciation of words sharing the same phonological components (i.e. the consistency effect), and with the phonological component itself (i.e. the congruency effect).

In a follow up experiment, Hue divided the congruent characters into congruent/consistent and congruent/inconsistent words. With all the congruent words sharing the same pronunciation with the phonological components, he was able to investigate the effect of consistency in addition to the congruency effect. If the inconsistent words were named slower than the consistent ones, the effect can only be attributed to the process of analogy - that is, deriving the pronunciation from words sharing the same phonological component.

Thirty-two Chinese participated in this experiment. Eight sets of words were selected, again crossing frequency (high and low) and word type (congruent/consistent, congruent/inconsistent, incongruent and unique words). Each set contained 10 words. The average frequency rank was 551 for high-frequency words, and 3865 for the low-frequency words. The procedure was the same as the first experiment reported above.

The results showed that high-frequency words were named faster than low-frequency words. The congruent/consistent words were named faster than unique words which in turn were named faster than congruent/inconsistent and irregular words. In addition, the interaction between frequency and word type was significant. There was no difference across word types for high-frequency words. However, for low-frequency words, congruent/consistent words were named faster than the other three types of words, unique words were named faster than congruent/inconsistent

words and no significant difference was found between irregular and congruent/inconsistent words.

The results of this experiment indicate that during the word recognition process, the phonological component can activate its own pronunciation, and, at the same time, the pronunciation of words with the same phonological components. The pronunciation of these activated words is used to synthesize the pronunciation of the target word. In the congruent/inconsistent and irregular conditions, subjects were more likely to make mistakes and took longer to name the words because they had to choose a pronunciation from a number of probable and possible pronunciations.

To sum up, adopting an analytical approach in investigating the phonological processing of Chinese allows us to conclude that the phonological component does, in fact, activate phonological processing during word recognition. The presence of the consistency effect and congruency effect have demonstrated that readers use analogy and derivation via phonological components for pronunciation. Not only the pronunciation of the phonological components but also the pronunciation of words sharing the same phonological components are retrieved in the process of naming a word. The phonological components provide facilitative, and at the same time, interference effects in reading because not all words with the same phonological components share the same pronunciation.

3.4 Orthographic aspects of reading Chinese

In the following section, some comparative studies on visual memory between the Chinese and alphabetic scripts are presented. All researchers agree that visual discrimination and memory are important to learn to read Chinese but the emphasis on its importance differs. Chen and Juola (1982) and Chen, Yung and Ng (1988) adopted a holistic approach, regarding individual words as logographs. They

believed the orthographic superiority of Chinese originated from the unique distinctiveness of the figural structure which facilitates visual memory. There is a large number of visually distinctive Chinese words and about 3,000 to 4,000 characters in frequent use. Readers have to relate each visual configuration to a sound and a meaning. The acquisition of this information is believed to be reached through processes dealing with the discrimination and identification of the appropriate word. With many thousands of different characters, the task is not trivial.

On the other hand, analytical studies (Huang & Wang, 1992; Fang & Wu, 1989; Lai & Huang, 1988) have looked into the orthographic structure of a character to explain the distinctive orthographic structure of Chinese which facilitates visual memory. Each Chinese word is a relatively complicated graphic unit but all of them are made up of strokes and stroke-patterns. The visual complexity of a word can be measured by the number of strokes in the words (e.g. Cheng, 1992; Seidenberg, 1985; Hue & Erickson, 1988) which vary from one to over 20. All the strokes are packed in a confined two-dimensional, square-shaped composition. The same does not hold for alphabetic scripts. The greater the number of letters, the longer the string. At the same time, visual complexity of words can also be measured by the number of stroke-patterns. Most of the compound words consist of two stroke-patterns only and the more complicated words consist of five to six stroke-patterns. They are familiar graphic patterns and most of the stroke-patterns occupy fixed positions although the sizes may differ in order to fit into the constricted squared-shaped composition (refer to Figure 2.2 in Chapter 2).

3.4.1 Holistic approach

Chen and Juola (1982) were interested to know if the orthographic information plays a more important role than the phonological information in reading Chinese, as compared with English. It has been suggested that the lack of symbol-to-sound correspondence encourages Chinese learners to apply a predominantly visual discrimination strategy. In this experiment, the roles of graphemic, phonemic and

semantic information on lexical coding and memory for Chinese logographs and English words were assessed by a memory task. It was predicted that if the processing strategies for Chinese and English were distinctive, then the information stored in memory should be coded in different formats. Specifically, if the dominant processing strategies for Chinese is visual coding and for English is phonological coding, then these memory codes should be differentially emphasized in retrieving information about Chinese words and English words.

They asked Chinese and English readers first to study a list of Chinese words or English words respectively, for a memory test later. The items on the list were selected from pairs of graphically, phonologically or semantically similar words. They found that the Chinese readers were faster and most accurate in matching graphically similar pairs in the recognition task; whereas there was little difference in the performance of the English readers. They concluded that visual representation plays a significant role in memory for Chinese logographs, whereas memory for English words resulted in a more integrated code involving visual, phonological and semantic information.

Chen, Yung and Ng (1988) also assessed the roles of graphic, phonological and semantic codes in language processing using a visual search task. The performance on the task presumably depends on operations that mediate the recognition of some features of the printed words in various contexts.

Twenty-eight native Cantonese-speaking adults were involved in the test. A visual search task was devised by showing subjects a string of nine items. Subjects were asked to judge if all the items in the string were similar in either orthographic, semantic or phonological aspects. The dissimilar string had one item that differed from the rest in the display along a specific orthographic, phonological and semantic feature. All the stimuli in this experiment were real words. The graphic words were similar because one of the stroke-patterns in the string of words was identical

although it was placed at different positions in the words. All the phonologically similar words were homophones, and the semantically similar words were synonyms. In each trial, a graphic / phonological / semantic clue was shown and ten seconds later, it was followed by a string of items. Subjects were asked to find out whether all items in the string contained that clued feature, or whether there was one, if any, which was different from the rest of the items in the string along a feature specified by the clue. A 'yes' response was required if there was one item that did not contain the clued information and a 'no' answer was required if all the items in the string contained the clued information.

The results obtained in this study, however, were quite contrary to Chen and Juola's (1982) study. The response time for deciding 'no' which meant that 'all the Chinese words in the string were the same' was slowest in the graphic strings and second slowest in the 'yes' responses when one item was different graphically from the others. This result seems unacceptable at first glance because the processing time in making a visual comparison should be faster than semantic processing and phonological processing where the words have to be transformed into sound or meaning. One possible explanation may be due to the confusing information provided by the components in the graphically similar string. All the words in the graphically similar string shared an identical graphic component. Apart from being graphic components, the authors had overlooked the fact that they were at the same time either semantic radicals or phonological components and they might provide semantic clues or phonological clues. This unwanted information obviously affected the recognition that the string of words were only graphically similar. The experiment has somehow demonstrated that Chinese readers do not regard a character to be the smallest perceptual unit. They use an orthographic unit smaller than a character to process orthographic, semantic and phonological information. Research adopting an analytical approach regards stroke-patterns, rather than the whole character as units for visual processing.

3.4.2 Analytical approach

Huang and Wang (1992) are convinced that a word is not processed as a Gestalt. They proposed that there are perceptual units within a word and they suggested a stage model to explain the visual process of Chinese word recognition. Non-accidental properties in the word such as collinearity, parallelism and cotermination are first searched for within the word and word regions are identified. Then, strokes and stroke-patterns and their relative location are activated. The activation of the stage model finally triggers the identification of the intended word.

The orthographic structure in a Chinese word consists of strokes and stroke-patterns. It is relatively easy to identify the strokes in a simple word (e.g. *sky* 〈天〉) and discriminate the two simple words with few strokes (e.g. *sky* 〈天〉 and *big* 〈大〉). However, it is not easy to discriminate the following three words by strokes 〈瑩〉 〈縈〉 〈榮〉 as they consist of 14 - 16 strokes. On the other hand, the three words look very similar because three out of the four stroke-patterns in these words are identical and they differ only from one another by one stroke-pattern, which is located at the bottom of the words 〈玉〉 〈糸〉 〈木〉. Thus if the words are processed visually by stroke-patterns instead of strokes, readers only need to focus on the differences and similarities of four stroke-patterns and the visual burden can be reduced significantly.

Fang and Wu (1989) hypothesized that strokes are the perceptual units in simple words and stroke-patterns are the perceptual units in compound words. A stroke is the smallest graphic unit in written Chinese and it is the smallest separable unit in simple words. A stroke-pattern is a cohesive perceptual unit and can be separated from other stroke-patterns visually as a whole in compound words. Thus Fang and Wu (1989) designed an experiment to examine the occurrence of illusory conjunction at two levels : the stroke level and the stroke-pattern level. Illusory

conjunction refers to the perceptual recombination of a component of a word with the other and it was predicted that with a short exposure of the target word and the distraction of a context composite word, a migration of a stroke or a stroke-pattern would take place.

Twenty native speakers of Chinese were asked to perform a target detection task. The twenty probe words and their corresponding test displays were evenly divided into the stroke-level conditions and the stroke-pattern level condition. Each trial consisted of the following sequence: a verbal "ready" signal given by the experimenter, a one-second presentation of a probe word, a one-second presentation of a black rectangle outlining the area within which the test words would appear, the test display for a variable duration, and finally a "checkerboard" noise mask for 200 ms. The test display consisted of two digits with three test words in between. The subject was instructed to report the digits and then to state whether or not the preceding probe exactly matched one of the test words.

In the three test items, there was a composite context word, and two source words. The composite context word was expected to enhance the separability of the source word components and thus create conjunction errors whenever illusory conjunctions were plausible. In the stroke-level condition, a stroke is expected to migrate to another word with the help of a composite context. For example, given two simple words *wood* 〈木〉 and *tongue* 〈舌〉 as source words and the composite context word *ancient* 〈古〉 with the probe word *grass* 〈禾〉, the slanting stroke on top of *grass* was expected to migrate to the composite context word *ancient* 〈古〉, and the word *tongue* 〈舌〉 would be identified as the probe word due to illusory conjunction. In the stroke-pattern condition, a stroke-pattern was expected to migrate to another word. The composite context word shared one stroke-pattern with each of the two source words. For example, the two source words were *wave* 〈波〉 and *service* 〈役〉, and the context word was *that* 〈彼〉 with the probe

word being *not* 〈没〉, the three dots from the probe was expected to migrate to the composite context character *that* 〈彼〉 and the word *wave* 〈波〉 would be identified as the probe word due to illusory conduction (Figure 2.6). Results indicated that there was no difference in the occurrence of illusory conjunctions at the stroke level in simple words and at the stroke-pattern level in compound words. There were actually more illusory conjunctions between the stroke-patterns than the strokes even though the difference was not statistically significant. Thus it is concluded that stroke-patterns are the primitive, separable units of word perception in over 90% of the compound words.

Figure 3.1 Probe words and test items for Huang and Wang's (1992) study

TABLE REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES

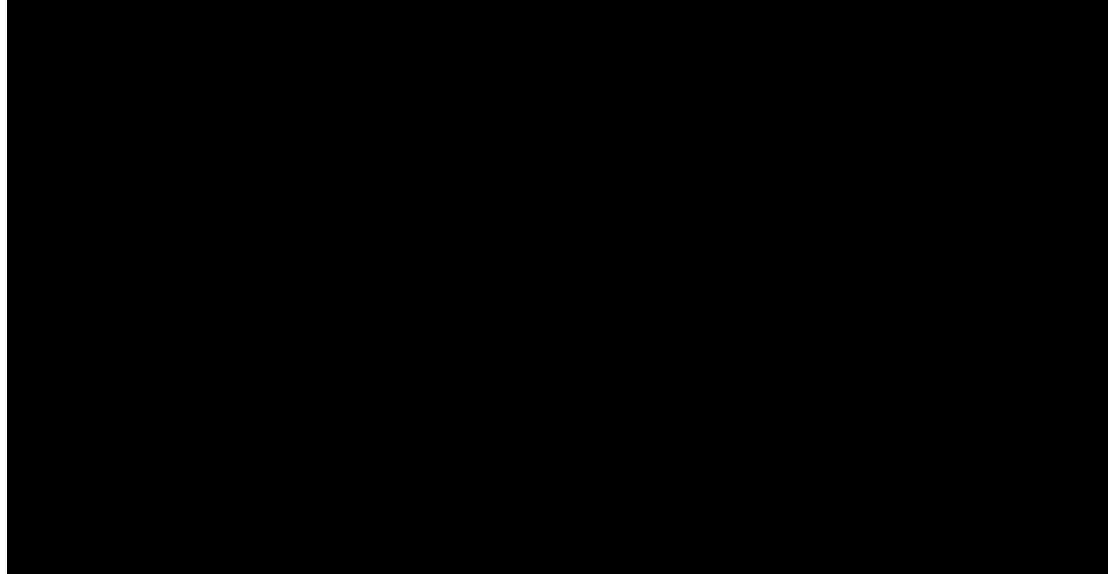
Chinese compound words are formed by combining different stroke-patterns on the basis of certain orthographic rules. The most significant orthographic rule is the positional regularity of the stroke-patterns. Lai and Huang (1988) investigated the pattern of component migration with reference to the stroke-patterns in Chinese words. They regarded the stroke-patterns as the graphic units of migration and traced whether there is location information in the formation of illusory words. They briefly presented horizontally-structured single-character-word pairs, AB and CD with each word composed of two stroke-patterns. A context word CB was

provided to facilitate the occurrence of illusory conjunctions (Figure 2.7). It was designed in a such a way that **AB**, **CD**, **CB**, **AD**, and **DA** were all legitimate Chinese words. The prediction was that the component **A** would remain at the left hand side and the component **D** would remain at the right hand side after the migration because that was their legal position in word **AB** and **CD**. Thus even though **DA** was also a real word, the percentage of forming a **DA** combination would be lower than a **AD** combination.

Thirty-two subjects participated in the task. The stimulus pairs were presented together with the context word for 50ms, then followed by a blank field for 50 ms. The target match was immediately presented and subjects were to tell whether **DA** or **AD** appeared in the stimulus pair. The result supported the prediction and the percentage of illusory words in the form of **AD** was significantly higher than that of **DA**. This means that the subjects were not migrating the stroke-patterns to any possible location, there is a strong positional bias accompanying the stroke-patterns during the visual component migration.

Figure 3.2 Sample material adopted in Lai and Huang's (1988) study

TABLE REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



A set of recurring stroke-patterns serves as the constituents of words. The same stroke-pattern, though recurring, could appear in different sizes at different locations in different words (e.g. 〈口〉 *mouth* in 〈狗〉 *dog*, 〈谷〉 *valley* and 〈唱〉 *sing*). Huang and Wang (1992) believed that unit frequency should be introduced in addition to Chinese character frequency count. Chinese character frequency count refers to the total number of characters appearing in isolation and unit frequency refers to the total frequency of stroke-patterns either by itself (as a simple character) or as part of a character (a stroke-pattern in a compound character). There are no statistical data available on the size of the set of recurring stroke-patterns but the frequency for most of the stroke-patterns is likely to be high. For example, unit frequency of the stroke-pattern 〈口〉 *mouth* is higher than 500. Thus the stroke-patterns become the building blocks for the orthographic structure for all the Chinese words, and visual skills are needed to discriminate the familiar stroke-patterns that may appear in different locations and in different sizes.

Even though the visual burden can be eased by the presence of the stroke-patterns, their positional regularity and recurring nature, it is emphasized that Chinese words are not memorized purely by visual memory. Research has demonstrated that very few Chinese words can be remembered purely by visual memory.

Zhang and Simon (1985) studied the short term memory span for semantic radicals, one-character words and two-character-words, which are regarded as being the basic language units in Chinese, using the immediate recall paradigm. The experiment had four goals: (1) to obtain a general idea of short-term memory spans for the most basic language units in Chinese (semantic radicals, characters, and words); (2) to probe the effect on short-term memory span of the presence and absence of pronounceable names for visual stimuli; (3) to test for the existence of separate acoustical and non acoustical short-term memory; and (4) to probe the effect of short-term memory span on the number of syllables in each stimulus item.

Three types of stimulus sequences were generated: a semantic radical sequence, consisting of a list of two to five semantic radicals; a character sequence, containing four to nine characters, and a two-character word sequence, containing three to six words. Each stimulus sequence was hand printed on a 35-mm slide with a blank space separating each pair of items. The slides were shown at a rate of about 750 msec/item. That is, if there were 3 semantic radicals on the slide, it would be exposed for $750 \times 3 = 2,250$ msec. Twelve native Chinese were asked to write down the lists of sequence in the right order. The response time was not limited. Memory span was computed from the number of items in the correctly recalled stimulus sequences.

Results showed that there is a significant difference in short-term memory span for semantic radicals, characters, and words. The mean short term memory span was 2.71 for semantic radicals, 6.38 for one-character words, and 3.83 for two-character words. Separate acoustical and nonacoustical short-term memory exist but the memory span for the latter was very small. All subjects considered it very difficult to recall semantic radical sequences because semantic radicals did not have common names. Thus the authors concluded that the reason why the memory spans for these semantic radicals are so small is because they did not have familiar pronounceable names.

The experiment has provided evidence that the memory span for purely visual recall is very small even among Chinese readers. The authors, however, did not control the character frequency effect and semantic radical frequency effect. The short-term memory capacity for high frequency characters and semantic radicals may differ from the low frequency ones.

Hue and Erickson's (1988) study also investigated immediate recall of Chinese semantic radicals and single-character-words, in an attempt to differentiate

contributions from visual and verbal-acoustic short-term memory. They improved Zhang and Simon's (1985) study by taking into account the linguistic frequency, inter-character frequency and stroke-complexity of semantic radicals; and in characters, the character frequency and stroke-complexity.

In one experiment, the immediate recall of individual Chinese semantic radicals was studied. In varying the frequency, orthographic complexity, inter-character frequency and linguistic frequency of the stimuli, the authors wanted to obtain a more comprehensive picture of memory span for Chinese semantic radicals. Semantic radical frequency was denoted as high if the semantic radical was pronounceable and low if it was not pronounceable. Orthographic complexity of semantic radicals was manipulated with high-complexity semantic radicals containing four or more strokes (mean = 6.92) and low-complexity semantic radicals containing three or fewer strokes (mean = 2.62). Linguistic frequency referred to the presence of the semantic radicals in the most commonly used characters. Those semantic radicals that were not included in the norm words were considered to be of low frequency. Inter-character frequency referred to the number of characters in which a semantic radical appears. High and low inter-character frequency semantic radicals were defined as those that appear in six or more and in fewer than six characters, respectively.

Ten Chinese graduate students participated in the study. Eight sets of semantic radicals were selected varying in linguistic frequency, inter-character frequency and complexity. Each set contained eight semantic radicals. The subjects attempted to memorize 64 lists, and the list length varied - some lists contained 2 semantic radicals, some lists contained 4 semantic radicals. The 64 lists were presented in random order at a rate of 500 msec per semantic radical; thus a card containing two semantic radicals was shown for 1 sec, a card containing four semantic radicals for 2 sec and so forth. The subjects were instructed to write down the semantic radicals after one list was presented.

The results showed that the memory span was higher for high-frequency semantic radicals than for low-frequency semantic radicals. The mean for high frequency semantic radicals was 4.8 and the mean for low frequency semantic radicals was 1.65. The interaction between frequency and inter-character frequency was significant. The low frequency semantic radicals were remembered better in high inter-character frequency than in low inter-character frequency. There was also an interaction effect between the frequency and stroke complexity. For high frequency semantic radicals, the complex semantic radicals were remembered better than simple semantic radicals; while the opposite occurred in the low frequency semantic radicals. The better recall of complex semantic radicals was a surprise to the authors and they gave one possible explanation, that is, complex semantic radicals are more distinctive. In fact, the phenomenon can be explained better by considering that readers are memorizing the distinctive stroke-patterns instead of strokes, as stroke-patterns can already be found in the high frequency complex radicals.

In another experiment, the authors used compound single-character-words as stimuli to examine short term memory by immediate recall. They expected a similar interaction of frequency and complexity to be obtained for compound Chinese single-character-words. High-frequency words were chosen from the 300 most frequent words and the medium-frequency from the next 3,000 most frequent words. The low-frequency words were chosen from words outside the 5,000 most common words. Care was taken not to select low-frequency words that contained a well-known phonological component. Six sets of words were selected by crossing frequency (high, medium or low) and complexity (high or low). Words with less than 6 strokes (mean = 3.4) were of low complexity and high complexity words consisted of 10 or more strokes (mean = 12.98). The procedure was the same as in the experiment mentioned above.

Eleven subjects were asked to remember 12 stimulus lists with eight items on each list. Each eight-item list was hand printed in random order on an index card, and

was presented for a 4-sec study period. The subjects were given as much time as they desired for immediate written free recall of each list.

The result showed that the high-frequency items were easier to recall than the low-frequency items. The mean for high frequency items was 6.4 and the mean for low frequency characters was 1.9. The interaction between frequency and complexity was also significant. For high- and medium-frequency words, effects of complexity were quite small, whereas, for low-frequency words, orthographically simple words were much easier to recall than complex ones. The presence of the stroke-patterns failed to facilitate the memory of less frequent words. The short exposure of stimuli (500 msec) did not allow the subjects to recognize all the stroke-patterns and reorientate the stroke-patterns to the appropriate locations. This assumption was supported by the visual errors made in the low-frequency words where only some stroke-patterns could be dictated.

On the whole, the comprehensive studies on short-term memory for Chinese semantic radicals and single-character-words has demonstrated that when words are relatively frequent and their pronunciations are well known, they seem to be maintained in verbal form in short-term memory. For those words of relatively low frequency, with pronunciations that are not likely to be known, they seem to be maintained in visual form in short-term memory. Memory span for visual short-term memory is much smaller than the verbal one. Subjects could only remember one to two words for low frequency words while they could remember about six words for high frequency words. Errors in dictating partial storage or recall of visual information occurred only on low frequency lists. The study has demonstrated that reading Chinese is more than a memory task for symbols.

3.5 Semantic aspects of reading Chinese

3.5.1 Holistic approach

Relatively few studies adopting a holistic approach have looked into the semantic aspects of reading Chinese. The majority of the Chinese characters are no longer logographs and it is almost impossible to guess the meaning of a character purely by its visual feature. Even for the very few still existing logographs, such as *fire* 〈火〉 and *half* 〈半〉, some explanation is needed in order to allow novel readers to get to the meaning. In addition, over 90% of the Chinese words use the semantic radicals to indicate the semantic categories of the words, and it is only in this component that the meaning of the word can be found. A holistic approach is inappropriate to investigate the semantic aspects of reading Chinese.

3.5.2 Analytical approach

Even though the meaning of a Chinese word is not conveyed through pictorial depiction, over 90% of the Chinese characters contain some semantic information in a graphic unit known as the semantic radical. In most cases, the semantic radical in a word indicates the semantic category to which the word belongs, although one is unlikely to obtain the exact meaning of the word solely on the basis of its semantic radical. In addition, there are irregular words in which the semantic radicals do not indicate the semantic category.

In Fei and Sun's (1988) study, the semantic relevance of semantic radicals was rated. The authors predicted that there was a difference in the degree of semantic information the semantic radicals could provide for readers to read different words. Adult readers were asked to rate 665 words that contained four semantic radicals based on the following indices: (a) a word whose meaning is directly related to the meaning of its semantic radical was rated as 10; (b) a word whose meaning has no relation to the meaning of its radical was rated as 0; (c) a word indirectly related to

the meaning of its semantic radicals was rated as 5. The results indicated that the four semantic radicals contributed to the word meaning differently, and the average degree of semantic radical usefulness varied from 2.5 to 6.9. Nevertheless, all four semantic radicals were helpful for deriving parts of the meaning of at least some words.

Flores d'Arcais (1992) investigated whether lexical decomposition is likely to take place during the recognition of a Chinese word to find out if the semantic information of a semantic radical could be activated when the word was recognized, even when the relation between the word and the radical was completely opaque.

Japanese subjects were required to make speedy judgments on the semantic relatedness of a pair of words: that is, to decide as quickly as possible if two words belonged to the same semantic category. If the two words were related in meaning, then the subject would quickly say 'yes', and when the words had no related meaning, the appropriate response was 'no'. There were four types of stimulus pairs: (1) graphically similar pairs (2) component-to-whole pairs where only the semantic component of a compound word was related to the meaning of the other word but the two words in the pair were in fact different in meaning. For example, the word pair *stone* 〈石〉 and *eye* 〈目〉 are different in meaning. The word *stone* 〈石〉 contained a component 〈口〉 and if the readers paid attention to this component, it might lead them to compare the semantic relatedness between the component *mouth* and the word *eye* as they belong to the same category for senses. (3) part-to-whole pairs where part of a word, that was a component in one word, was the other whole word in the pair, for example, the pair *stone* 〈石〉 and *mouth* 〈口〉. (4) Unrelated pairs where words were neither similar nor semantically related.

Results from this study showed that there were significant differences in response latencies and error rates between the four types of stimulus pairs. The component-to-whole pair were judged slower than the part-to-whole pairs, and this result was interpreted as being due to interference in the decision of semantic relatedness for the pairs consisting of an opaque semantic component (*eye* and *stone* contain as an opaque component the radical *mouth*). The subjects were drawn to a 'yes' response which interfered with the execution of the correct response 'no'. In other words, the presentation of the word *stone* would activate not only the correspondence representation in the mental lexicon, but also the component *mouth*.

It is interesting to see how the 'opaque semantic component' can be activated by some graphic components in Chinese words. The result also supports the notion of some form of lexical decomposition process in the recognition of the compound words in adult readers. However, it is impossible from this experiment to conclude that the semantic information corresponding to a (semantic) radical becomes available and is functionally effective when the compound word is recognized. The author was unable to define the term "radical" clearly. He regarded *mouth* as a radical of the word *stone* and *eye*, but *mouth* is not a semantic radical of *stone* and *eye* in the Chinese dictionary. In the Chinese dictionary, both words *stone* and *eye* are semantic radicals themselves. The word *stone* is not a compound word either. The semantic components in a compound word need to be well-defined in order to investigate the function of the semantic radicals.

Zhang, Zhang & Peng, (1990) studied the role of semantic radicals in retrieving word meanings. The question asked in the study was whether the semantic relevance of the semantic radicals would affect the decision time in a semantic comparison task. Subjects were exposed to a target word and a distracter, and were asked to decide whether they were related to a semantic category. Two types of words were investigated. The first type consisted of words with a semantic radical consistent with the semantic category to which the word belonged. For example, the

word *aunt* 〈女〉 + 〈古〉 = 〈姑〉 with a female radical 〈女〉 semantically belongs to the female category. The distracter was a word without the female radical but belonging to the same category, such as the word *mother* 〈母〉 (there are several ways to write *mother*, such as 母, 媽, 娘). The second type consisted of words without the category-indicating semantic radical (e.g. the word *brother* 〈哥〉), and the distracter here was a word with an irrelevant semantic radical, but in fact, semantically belonging to the same category as the target word, such as the word *son-in-law* 〈女〉 + 〈胥〉 = 〈婿〉. Reaction time for the type of words with relevant semantic radicals was faster than the second type of words without category-indicating semantic radicals. The study thus indicated that retrieval of word meanings is facilitated when the semantic radical of a word is consistent with the category of the word and inhibited when it is inconsistent.

In another study, Flores d'Arcais (1992) investigated the role of semantic radicals in retrieving word meanings in a different setting. The task was set up in such a way that the radical was shown first and then followed by a whole compound word. It was predicted that by pre-exposing a radical which was congruent in form and in meaning with the word, a preactivation of the meaning of the word could be obtained and would result in a faster processing time. On the other hand, an inhibition of the word would be observed if the radical was incongruent in meaning with the radical of the compound word. "Radical" was defined by its positional regularity, all the components on the left side of words were treated as (semantic) radicals.

Sixty-four adult Chinese participated in this study. The basic material consisted of 20 pairs of compound words. The radicals in the pair of words were identical. In one word, the radical was related to the meaning of the whole word, and in the other word, the meaning was not related. For example, with the radical *canoe* 〈舟〉,

the congruent word was *boat* 〈舟〉 + 〈沿〉 = 〈船〉 and the incongruent word was *manner* 〈舟〉 + 〈役〉 = 〈般〉. In the control condition, a word irrelevant in meaning and orthographic form was presented. Each word was presented in isolation and had to be named as quickly as possible. The different onset times between the pre-exposed component and the whole word were 60 and 180 msec.

The result showed that with a short Stimulus Onset Asynchrony (SOA), the radical had a facilitatory effect both for a congruent meaning and for an incongruent meaning. The processing of both were faster than the control. This indicated that earlier exposure of a radical which was congruent in form (orthographically identically) had facilitatory effects, independently of the semantic relationship. At longer SOA, the semantically congruent radical facilitated the naming of the whole word which was related in meaning to the radical, but when the pre-exposed radical was incongruent in meaning with the name of the word, there was some inhibitory effect.

The result was interpreted as indicating an early, initial activation of graphic information concerning the form of the radical. This produces facilitation of the recognition of all words which share that form, independent of the meaning relation. Subsequently, with a longer SOA, there is an effect of activation of the semantic information. The pre-exposed radical has a given meaning, and all, or perhaps the most frequent words which have a related meaning are preactivated. When the target item is inconsistent in meaning with this cohort, inhibition may result.

The activation processes of the orthographic and semantic information have been demonstrated in this experiment. However, the definition of a 'radical' in this study is again incorrect. Flores d'Arcais defined the radicals simply by their positional regularity, that is, the left hand side of a word. It is true that in a lot of Chinese

characters, the semantic radicals are on the left hand side, but at the same time, semantic radicals can appear on the right, on top or at the bottom of the word. Neglecting the semantic radicals that can appear in other positions is misleading and makes generalization of the results impossible.

3.6 Conclusion on skills of fluent Chinese readers

There are two ways in which adult readers adopted in reading Chinese: the holistic approach and the analytical approach. The holistic approach regards a character as a logograph. The graphic, semantic and phonological information are believed to be processed by stroke-complexity, synonym effect and homophone effect respectively.

An analytical approach to the reading process assumes that three kinds of stimulus information facilitate reading: (1) information dealing with the figural structure of the word; (2) information concerning the pronunciation of the word and (3) information concerning meaning. There is evidence that stroke-patterns play a crucial role in analyzing the complicated figural structure of Chinese words, and adult readers are aware of the positioning of stroke-patterns. Quite a number of studies (Seidenberg, 1985; Zhu, 1987; Fang, Horng & Tzeng, 1986; Hue, 1992) looked into the importance of the phonological components in pronunciation. Words can be read via the pronunciation among different words with the same phonological component (phonological consistency) or directly from the phonological component (phonological congruency). Recently Flores d'Arcais (1992) and Zhang et al (1990) have also demonstrated that semantic information in the semantic radicals is utilized by adult readers to get to the meaning. However, the above-mentioned experiments were hampered by some serious drawbacks in that they either failed to define the semantic radicals and the phonological components in a word clearly or they have restricted their stimuli to one type of ideophonetic words which account for a mere 39% of all written words.

Although the orthographic structure of Chinese and that of alphabetic scripts is very different, the processing of various kinds of information is quite similar. For example, it has been demonstrated in both English and Chinese that high frequency words are maintained in verbal form and low frequency words are maintained in visual form (Zhang and Simon, 1985; Hue and Erickson, 1988). It is, therefore, believed there is a general principle of fluent reading that applies to all writing systems.

3.7 Using analytical approach to investigate the processing of orthographic, semantic and phonological information in reading Chinese

It is only recently that researchers have taken up a theoretical position in investigating the various linguistic information provided by the stroke-patterns. Flores d' Arcais' (1992, 1995) and Chen's (1993, 1995) believe that it is stroke-pattern, rather than character that provides linguistic information in reading Chinese.

3.7.1 Flores d'Arcais' (1992, 1995) study

Flores d'Arcais (1992) has set the scene for lexical decomposition and has provided evidence of the activation of orthographic, semantic and phonological information in reading Chinese. Assuming that semantic and phonological information are encoded in different arrays in Chinese, the author investigated the phonological and semantic activation processes in the recognition of Chinese words in a naming task. The presentation was characterized by an onset asynchrony between part of the word and the whole, in such a way that a component (a component on the left is a radical and a component on the right is a phonological unit) either on the left or the right was presented before the whole word, the remaining part being withheld for a short interval. By momentarily withholding part of the information given by the word, the author investigated the time course of activation of the phonological information as compared with the semantic one.

Twenty-four Japanese were involved in the study. Four sets of 10 compound kanji, with radicals on the left and phonological components on the right were adopted. The four conditions of stimuli were as follows: in condition (a) the left radical was related in meaning and the right radical had an identical pronunciation as the whole word (sem+phon+); in condition (b) the left radical was related in meaning but the right radical had a different pronunciation from the whole word (sem+phon-); in condition (c) the left radical was unrelated in meaning to the whole word and the pronunciation of the right radical was identical to that of the whole word (sem-phon+) and finally, in condition (d), the whole word differed in meaning from the meaning of the left radical and in pronunciation from the pronunciation of the right radical (sem-phon-). Subjects were required to name the words as quickly as possible.

The results showed that at a short Stimulus Onset ASOA (60 msec) between the onset of the radical and the onset of the whole word, the prior exposure of the semantic radical facilitates access to the name of the whole word, and so does anticipation of the phonological component. The facilitative effect may be caused by orthographic activation of components with the short exposure of radicals. At 180 SOA, there is a significant effect from the prior exposure of the phonological component. This has a strong facilitatory effect on the naming of the word. On the other hand, at this SOA, the previous exposure of the semantic radical seems to have some inhibitory effect on naming speed. This study was replicated recently (Flores d'Arcais, Saito, & Kawakamu, 1995) with a larger sample of Japanese undergraduate students and same results were found.

However, there are some shortcomings in this study. Because the naming task is more likely to activate phonological information rather than semantic information, a more appropriate task is needed to detect the interrelationship between the orthographic, semantic and phonological information. An obvious flaw in this experiment is again the definition of the radicals. Flores d'Arcais used the term

'right radical' and 'left radical' to represent the semantic radicals and the phonological components. This is, in fact, incorrect, because the author has neglected a large number of words with the semantic radicals and phonological components not in the left-right position. A more precise definition of the orthographic units is needed to allow further investigation in the process of reading Chinese.

3.7.2 Chen's (1993, 1995) studies

Chen (1993) has contributed a comprehensive piece of work looking into how adult readers processed orthographic, phonological and semantic information at an intra-character level in reading and traces how the different kinds of information are interrelated. Her study made a clear distinction between the orthographic units in Chinese words in terms of their structure, independent of etymological origin. The orthographic rules of lexical decomposition were introduced to locate the semantic radical and phonological component. The structure of the two unique orthographic constituents - the semantic radical and the phonological component - was clearly defined. She has also demonstrated that phonological recoding rules of 'derivation' and 'analogy' are adopted by adult skilled readers. I would like to give a detail description of Chen's theoretical framework and discuss her findings in more depth, as my study will follow a similar structural approach, so as to explore whether the orthographic and phonological rules applied by adult skilled readers are also adopted by children when they begin to learn to read.

Eleven experiments were carried out which included visual comparison tasks, phonological comparison tasks and semantic comparison tasks. Twelve native readers, aged 25 to 35, who are fluent in Chinese were invited to take part in all the tasks.

3.7.2.1. The orthographic unit of Chinese words - the smallest stroke-pattern

In explaining the visual analysis of over 90% of Chinese words, Chen, Allport & Marshall (in press) proposed stroke-patterns, rather than strokes, as the functional orthographic units of Chinese words. In one of her experiments, Chen attempted to find out whether strokes or stroke-patterns are used by Chinese adult readers in visual analysis of words. Her rationale was that if strokes are the functional units of Chinese words, then speed and / or accuracy of visual comparison should be affected by the number of strokes (stroke-complexity). In the same way, if stroke-patterns function as the constituent units of Chinese words, then the speed and/or accuracy of visual comparison should be affected by the number of stroke-patterns (unit-complexity) of the "same" pairs and/or the number of mismatching stroke-patterns in the "different" pairs (unit-dissimilarity).

Twelve Chinese adult readers were asked to state whether pairs of words, visually presented on a computer, were the "same" or "different" in terms of their visual form. There were three word types (real words, pseudowords and nonwords); two levels of unit-complexity (two and three units of stroke-patterns in a word); and two levels of stroke-complexity (5 - 7 strokes in simple words vs. 8 - 11 in complex words). "Different" word-pairs also varied in the number of mismatching stroke-patterns in a word (unit-dissimilarity). There were five levels of unit-dissimilarity: same, 1/3 same, 1/2 same, 2/3 same and different. Pseudowords were constructed by substituting one unit from a real word to form a non-existent combination of orthographic units. Semantic radicals of pseudowords were always in their legitimate positions. Nonwords were constructed by placing the semantic radicals of the pseudowords in illegal positions (e.g. 永 is a nonword because the legal position of the three dots should be on the left rather than on the right.)

A highly significant word-type effect was found. Reaction time (RT) increased markedly from words to pseudowords to nonwords for both "same" and "different"

judgments. For "same" judgments, it was found that the difference in RT between the two-unit and three-unit stimuli (i.e. unit-complexity effect) was significant for both words and pseudowords, but not for nonwords. The unit-complexity effect remained significant when the number of strokes was controlled as there was no significant stroke-complexity effect on RT in any of the conditions. Furthermore, for "different" judgments, RT was significantly faster for pairs of words with a greater number of mismatching stroke-patterns. The effect was evident for words and pseudowords, but not for nonwords. In general, the pattern of error rates was similar to that of RT.

Chen concluded that the significant unit-complexity and unit-dissimilarity effects and the non-significant stroke-complexity effect suggest that the smallest stroke-pattern, rather than individual stroke, is the functional orthographic unit in the visual analysis of Chinese words.

Furthermore, the effects of both unit-complexity and of unit-dissimilarity were most evident for word-comparisons, somewhat less in pseudoword-comparisons and in nonwords. This suggests that recognition of individual orthographic units is better in a word context. Moreover, comparisons of word-pairs were significantly faster than on pseudowords, and pseudowords in turn were faster than nonwords. The effect of superiority of words over pseudowords indicates the functional role of lexicality (and perhaps visual familiarity) in the visual analysis of Chinese words. The advantage of pseudowords over nonwords indicates that positional regularity of the semantic radicals may play an important role in the visual analysis of Chinese words. These two types of word-superiority effect suggest that there is mutual interaction between orthographic units and words in the visual processing of Chinese. In general, these results suggest some functional parallelism between orthographic units in Chinese and letters in English. Such orthographic units as the smallest stroke-pattern can be regarded as functional orthographic units in Chinese. Stroke-patterns create a level of orthographic units intermediate between words and

individual strokes. In addition, they are also linguistic units as they have full lexical status, that is, they are both pronounceable and meaningful.

With regard to the visual comparison of strokes and stroke-patterns, Chen has assumed that they are the only two levels of graphic units. In fact, strokes can also create some kinds of graphic units which are not stroke-patterns at all. Therefore, there are actually three levels of graphic units: strokes, graphic units and stroke-patterns. Chen's study has not demonstrated whether the readers regarded the stroke-patterns as graphic units, or as the 'proper' stroke-patterns exclusive to Chinese which are not only limited in number but also in shape.

3.7.2.2. Two functionally distinct orthographic constituents of compound words : the semantic radical and the phonological component

Chen and Allport (1995) proposed that 90% or more of the Chinese words, can be divided into two principal orthographic types - the semantic radical and the remainder : the phonological component. The semantic radical refers to the 189 indices listed in the Chinese dictionary when these components occur in the radical position in a word. The semantic radicals can only be defined in word context, not in isolation. The relationship between the semantic radical and the whole word is to be called the rule of Part-to-Whole semantic correspondence because the meaning of part of a word (the semantic component) can be referred to as the meaning of the whole word. The phonological component is defined simply as all the remainder of the word, excluding the semantic radical. The relationship between the phonological component and the whole word is to be called the rule of Part-to-Whole phonological correspondence because the phonological information of the whole word can be found in part of the word (the phonological component).

In order to investigate whether the rule of Part-to-Whole semantic correspondence and the rule of Part-to-Whole phonological correspondence are utilized by adult

readers in word recognition, a series of phonological comparison and semantic comparison tasks was carried out.

In the phonological comparison task, Chen predicted that the physical matching of one component may function as a distracter and would bias subjects toward false "same" responses. If subjects had a selective attentional bias towards the phonological component for phonological comparison, then the word-phonological component match condition should elicit more false "same" responses than the word-semantic radical match condition. In addition, it is also believed that phonologically congruent and consistent word pairs would elicit more false "same" responses than the incongruent and inconsistent pairs.

Twelve Chinese adult skilled readers were asked to judge whether pairs of words had the same or different pronunciations. All the word-pairs were presented visually on a computer. There were several word-pair conditions. Half of the pairs were pronounced the same, and half were different. In two of the different-pair conditions, the pairs consisted of one simple word and one compound word in which the simple word was embedded as a component. In the word-semantic radical match condition, the simple word was embedded in its paired compound word as the semantic radical (e.g. *woman* 女[nǚ] - *good* 好[hǎo] where 女 is the semantic radical of 好). In the word-phonological match condition, the simple word was embedded in its paired compound word as the phonological component (e.g. *stage* 台[tái] - *joy* 怡[yí] where 台 is the phonological component of 怡). These word-pairs were neither homophones nor synonyms. The pairs in the third different-pair condition consisted of two compound words with identical phonological component (e.g. *tear* 揭[jiē] - *drink* 喝[hē] where 曷 is the phonological component of both words).

For the same-pair condition, there were two types of stimuli: (1) the pairs of words consisted of one simple word and one compound word in which the simple word was embedded as the phonological component (e.g. *warehouse* 庫 [kù] - *trousers* 褲 [kù] where 庫 is the phonological component of 褲) and (2) the pairs of compound words with identical phonological components (e.g. *smelt* 煉 [liàn] and *practice* 練 [liàn] where 煉 is the phonological component). This time, the word pairs were homophones. Word pairs with no physical similarity served as the baseline.

Chen also calculated the congruency and the consistency index for her stimuli. The congruency index refers to the degree to which pronunciation of the compound words is congruent with the pronunciation of its phonological component in isolation. The consistency index is an index of the degree to which compound words with the same phonological components are consistent in pronunciation with each other.

It was found that the mean reaction time for the word-phonological match was significantly slower than that for the word-semantic radical match in the phonological task. The error rates (i.e. the rate of false "same" responses) for the word-phonological match condition was significantly higher than for the word-semantic radical match. The word-phonological match elicited higher error rates than the baseline but the word-semantic radical match did not. Furthermore, the congruency and consistency indices were found to be significantly and positively correlated with the reaction time and error rates of the word-semantic radical and word-phonological match pairs. In other words, the more phonologically congruent and consistent the compound word was, the more false "same" responses it generated and the longer the time needed for sound comparison. In a series of multiple regression analyses, only the congruency index remained significant. Chen's (1993) study has demonstrated, therefore, that in skilled readers of Chinese

there is a selective attentional bias toward the phonological component in a phonological judgment task. These findings demonstrate that the phonological component in compound words plays a significant role in adult readers' pronunciation of Chinese words, and provide some support for the psychological reality of the "orthographic-phonology correspondence" rules in skilled reading of Chinese.

Derivation is believed to be the major procedure for pronouncing Chinese compound words, following lexical decomposition. The congruency index was significantly associated with the comparison which indicates that lexical decomposition and phonological recoding in Chinese may be based mainly on the phonological congruency between a word and its own non-radical component.

The same set of stimuli was used again in semantic comparison tasks of pairs of Chinese words which were conducted to examine the functional role of the semantic radical, in comparison with that of the phonological component, in the semantic recoding of Chinese words. The semantic comparison experiment was designed in parallel with the phonological comparison task mentioned above. Chen asked the same group of subjects to judge the same set of materials as to whether the pair of words had the same or different meanings. Test procedures were the same as those in the phonological comparison task.

It was found that the word-semantic radical match pairs elicited significantly more false "same" responses than the word-phonological match pairs for semantic comparison. The semantic comparison of the word-pairs was significantly affected by the semantic radical as a distracter toward false "same" responses, but not by the phonological component (word-component match). This is precisely the opposite result found in the phonological comparison task with the same stimuli and the same subjects. Results from the experiments suggest that subjects have a selective

attention bias toward the semantic radical in semantic judgments and towards the phonological component in the phonological judgments. That is, the semantic radical is likely to receive priority of attention in semantic tasks, while the phonological component may become salient in phonological tasks.

3.7.2.3. Orthographic-phonology correspondence rules : Derivation and Analogy

Another finding of her study is that the pronunciation of over 90% of Chinese words may be achieved via a rule-governed lexical decomposition process. The process of phonological recoding via orthographic components in Chinese can be achieved by either (1) by derivation, based on phonological congruency between a word and its own phonological component in isolation or (2) by analogy, based on consistency among words with the same phonological component. Chen called these "orthographic-phonology correspondence" rules.

In one experiment, Chen examined how adult readers apply the orthography-phonology correspondence rules to read pseudowords. Although a pseudoword virtually has no correct pronunciation in Chinese, it was predicted that subjects would guess the pronunciation of pseudocompounds using derivation, and guess the pronunciation of the pseudosimples by analogy.

Thirty pairs of pseudoword-word pairs was chosen as stimuli, and they were divided into two conditions: (1) half of the pairs consisted of pseudosimples created to match with a real word. The unpronounceable pseudosimple appeared by itself and also as a stroke-pattern in the word stimulus; (2) half of the pairs consisted of a word and a pseudocompound. The word became a stroke-pattern in the pseudocompound. In both conditions, half of them were semantic radical-word match and the other half were phonological component-word match. Subjects were

required to guess whether the pair of words had the "same" or "different" pronunciations.

It was demonstrated that subjects were able to use analogy and derivation to pronounce the pseudowords. In addition, the 'same' response rates toward the phonological component-word match were significantly higher than in the semantic radical-word match for both the pseudosimples and pseudocompounds. Pronunciation of the 'pseudosimple', achieved by analogy from the real word containing that same stroke-pattern, is more likely to be elicited with a shared phonological component, rather than a shared radical component. This further supports the idea that skilled Chinese readers may have a selective attentional bias toward the phonological component in a phonological judgment task, irrespective of the lexical status of the graphic component.

Although significant results were found in Chen's study, it is reminded that she had simplified the Chinese word recognition process. First, Chen has only considered the compound words. And second, the experiment failed to draw a definite conclusion if readers are using analogy or derivation in reading, as the reading responses alone cannot review the entire reading process that is activated by the readers. Simply reading a whole character as the phonological component does not necessarily mean that the reader is using derivation in pronunciation. Third, Chen does not take into consideration the semantic information that can be conveyed by the phonological components, such as indeocompound words. Another drawback in her study was her choice of subjects. There were only twelve subjects in the study, and the same subjects were used in all the experiments, it was not sure if familiarity to the experimental setting affected their performance.

3.7.2.4. Conclusions of Chen's (1993,1995) studies

To sum up, Chen's (1993, 1995) study has shown that Chinese so-called logographic words can be read analytically by adult readers on the basis of orthographic units smaller than a word. Instead of a character, or a stroke, Chen proposed that the stroke-patterns should be treated as the basic orthographic units because they have full lexical status, that is, they are both pronounceable and meaningful. In her study, she has demonstrated that stroke-patterns are utilized by skilled readers of Chinese in reading Chinese words, although the subjects have not been taught explicitly about such segmentation of orthographic units.

Chen's study has also demonstrated how adult readers elicited semantic and phonological information by means of lexical decomposition in order to pronounce and understand the meaning of the words. A framework is introduced in which 90% of the Chinese words can be systematically divided into semantic radicals and phonological components. In addition, adult readers follow the rules of 'derivation' and 'analogy' to derive the pronunciation of the word.

An important point to note here is that this systematic process in arriving at the meaning and sound of words is again implicitly learned by adults. This set of skills has not been taught in school although fluent readers manage to acquire them gradually throughout the learning process. At the same time, there is reason to believe that this set of skills is needed right at the beginning of learning to read. Young children are exposed to a lot of new words to be learned every day and since it has been proved that very few words can be remembered purely by rote memory (Zhang & Simon 1985; Hue & Erickson 1988), children have to construct some rules in order to learn to read. It is predicted that children at some stage develop the analytical skills acquired by the adult readers. For example, it would be of great help in the process of learning to read if they were able to acquire the rules to understand the orthographic structure of written Chinese, knowing that words are

formed by combination of a limited set of stroke-patterns. There are legal positions for these stroke-patterns and violation of these orthographic rules will lead to the formation of nonwords. At the same time, the ability to decompose the words into semantic radicals and phonological components would help children to get to the meaning and pronunciation of new words by themselves. In the following chapter, research that has been carried out on children learning to read Chinese will be reviewed in order to answer the following questions: (1) What is the developmental process of learning to read Chinese? (2) At what age do children acquire similar skills in reading which are comparable with the skills of adults? The pedagogy of teaching reading in Hong Kong, China and Taiwan will also be described in order to illustrate how children are expected to learn to read Chinese in school.

Chapter 4

Children learning to read Chinese

4.1 Introduction

In this chapter, research which has been carried out to investigate the process of children learning to read Chinese is reviewed under two headings: the pre-reading stage and the formal reading stage. In the first section research on Chinese preschool children will be reviewed. Research on preschool children has provided evidence to demonstrate how young children are actively involved in literacy activities before they are formally taught to read and write. Some development in their understanding of the orthographic, phonological and semantic aspects of written Chinese can be traced at this stage. Most of the children go to primary school with some kind of knowledge about the written system. At six, children are taught to read and write formally. There is only a handful of studies to investigate the process of children learning to read. Of these, nearly all of them have focused on the relationship between visual skills, phonological skills and reading ability. The second section of the chapter discusses the contributions of the existing research to our understanding of children's development of the orthographic, semantic and phonological aspects of written Chinese. Before going on to the literature review, a brief introduction to the teaching of Chinese in the primary school years helps to set the scene showing how educators think, or how they expect children learn to read Chinese.

4.2 The pedagogic perspective on teaching written Chinese

4.2.1 Reading difficulties of Chinese children

Chinese is the first language for children in Hong Kong, Taiwan, and China. Some time in the 1970s, it was suggested that Chinese, which is logographic in nature, poses

no difficulty in learning to read; however, that is not a realistic stance. Rozin, Poritsky & Soetky (1971) claimed that Chinese symbols are easier to differentiate than English words because of their configurational distinctiveness. On the other hand, Chinese educators, in general, believe that children have difficulty in learning to read because Chinese words are numerous and complicated. Stevenson and psychologists from Japan and Taiwan conducted an extensive study comparing the reading abilities of children in the fifth grade with those in their own country (Stevenson et al, 1984; 1985). They found that the overall incidence of reading backwardness among children in Japan, in Chinese-speaking countries (in this case Taiwan) and in the West (the U.S.A.) is very similar. There are children in all three countries who are reading at least two grade levels below their own grade.

There is evidence of individual differences in learning to write, which implies that not all children progress at the same rate in literacy development. Yang (1990) studied the written language ability of Chinese first-grade children in Taiwan. One hundred Chinese children were asked to write a spontaneous picture story and to copy ten words. The results indicated that there was a wide range of performance amongst the first-grade children in their productivity, use of syntax and vocabulary. For instance, in terms of quantity, four children were unable to write more than ten words whereas five children wrote more than 100 words. In analyzing the sentence patterns, the children used mainly simple sentences in describing the picture. About 15% of them wrote compound and complex sentences, but there were four children who could not make one complete sentence. Significant differences were also observed in the choice of word types and use of vocabulary. The high achieving children wrote more types of words and used more descriptive adjectives than the low achievers. The findings of this study are alarming because the study was conducted in the first semester of the first year, only a few months after the children have been formally introduced to reading and writing. Nevertheless, some children had already fallen behind in almost every aspect of their written language proficiency.

4.2.2 Methods of teaching Chinese

Assuming that learning Chinese words is a heavy burden for young children, Mandarin Phonemic symbols (Zhu-yin-fu-hao) were introduced in Taiwan, and Pin-Yin (a phonetic system) was adopted in China. First graders are introduced to the phonetic systems before they are formally introduced to Chinese characters. The rationale for introducing phonetic symbols is to relate children's knowledge about oral language to their learning of written language through a more consistent sound system. Dai and Lu (1985) reported a teaching experiment in the use of Pin-Yin as a preliminary stage in teaching reading and writing in mainland China. The experiment was designed to improve Chinese language teaching, and was instituted in the first year of school. The main stages of the experiment were as follows: on entering primary school, pupils were first taught Pin-Yin until they were able to read words very quickly in Pin-Yin. They then read a large number of parallel texts, in which Pin-Yin was placed side by side with Chinese characters. At an earlier stage than usual, they then began to write sentences using Pin-Yin and the words they had learned. After the first year, an 'Experiment Report' was published and it was found that the reading and writing ability in the experimental class receiving intensive Pin-Yin drills was comparable to the third year children who had not received any training in Pin-Yin at all.

Lee (1989) challenges the value of teaching phonetic symbols in the beginning reading program as she has gathered evidence that children can learn to read Chinese even before Pin-Yin is taught to them. Moreover, assuming that knowledge of Pin-Yin can help children to relate written words to spoken language, it offers little help to enable children to memorize individual words and to recognize words if Pin-Yin is not provided. Thus the problem of decoding words in isolation without Pin-yin has not been solved.

In most Hong Kong schools, Pin-Yin is not taught, and children are required to learn Chinese characters directly. Emphasis in teaching is placed initially on the

recognition of words, and then proceeds to sentences and passages. Most of the time in the early primary years is spent in teaching new vocabulary, while reading and composition are postponed to a much later stage. Children are only encouraged to read independently at the age of ten, and for writing, children are only expected to write about 100 words at the age of twelve.

Word decoding skills are believed to be crucial in learning Chinese but for thousands of years, rote memory has been the only method used by educators in teaching vocabulary. Children are required to read and write numerous words everyday, and dictation, tests and examinations are used to put extra pressure on children to work diligently at their language studies.

Recently Fan et al (1987) have proposed a new teaching method called the concentrated method of learning Chinese words. The main characteristic of this teaching method is to teach the rules for word structure which can facilitate children's learning. This method stresses the structural rules of ideophonetic compound words. It puts together the words that have the same pronunciation and share the phonological components. In this way, children can draw inferences about other cases from one instance, and can learn words much faster.

Before any fundamental change in teaching methods can take place, it would be advisable to develop more knowledge of how children learn to read Chinese in order to provide valuable information for the design of a suitable teaching approach.

4.3 Literacy development at pre-reading stage

Chan's (1990) study and Lee's (1989) study traced the active contribution of children to their own learning of reading and writing before formal schooling. They assumed

that children at a very early age show interest in the script in their environment. Ho (1994) investigated the precursors of learning to read Chinese through a two-and-a-half year longitudinal study when the children were 3 years 4 months old. She focused on pre-school children in three studies and the findings give us an idea about the initial stages of literacy development among Chinese children. The results of these three studies will be reviewed under the orthographic, semantic and phonological aspects of written Chinese.

4.3.1 Orthographic aspects of written Chinese

4.3.1.1 Developmental stages in understanding the orthographic structure of written Chinese

Chan (1990) investigated preschool children's understanding of written Chinese and traced the developmental changes in their knowledge of the writing system. The study followed Ferreiro's work (Ferreiro, 1978; 1985; Ferreiro & Teberosky, 1982) in describing the developmental stages of written language acquisition, and contributed to an understanding the nature of children's hypotheses about the written language.

Sixty children, aged 2 to 6 years were invited to complete the following task: 1) draw a picture of self and write their own names; 2) write a one-character-word *house* 〈屋〉, a two-character-word *mother* 〈媽媽〉, and a three-character-word *elephant* 〈大象〉.

Results from the study indicated a clear developmental trend in the understanding of the orthographic structure of written Chinese by preschool children. Figure 4.1 presents some examples of children's writing of house, mother, elephant, and their

own names in various age groups. Their writing can be allocated to the following developmental stages:

Stage one: Differentiation between drawing and writing

Three years old children do not share with adults the knowledge that print is 'language written down'. They do not suppose that print represents language, instead they move back and forth between picture and text. Kay (3 yrs 1 mth) drew a house for the word *house*, a human figure for the word *mother* and an animal for the word *elephant*. But when she wrote her own name, she scribbled instead of drawing another figure. Ferreiro and Teberosky (1982) noted that name writing is 'the first stable written string and the prototype of all subsequent writing'. Kay had acquired some rudimentary concept of print in her name writing by using a 'writing' mode. Chan (3 yrs 8 mth) was more advanced in his concept of print than Kay. He had already developed the idea that writing is different from drawing, he gave up pictorial depiction and used abstract drawings which look more like writing. For example, he drew only the nose of the elephant instead of the whole animal believing people could guess what he meant by the long nose of the elephant. His name writing was again at a more advanced stage than his writing of the other words, as he could create a stroke-pattern which already looked like his surname 〈陳〉. Thus in his name writing, he had entered a further stage of development.

Stage two: Establishment of standardized graphic patterns

At the second stage, children have to attend to certain graphic properties of written Chinese. Lo (4 yrs 8 mth) had completely given up drawing in his writing. He used simple strokes, such as vertical and horizontal strokes, and dots in his writing. Two stroke-patterns were also found, the stroke-pattern *mouth* 〈口〉 and another invented stroke-pattern for his own name. He also used an inappropriate stroke-pattern - a circle for the word *elephant* - although the circle is never used as a stroke-pattern in Chinese writing. At this stage, children have to figure out the kind of strokes and stroke-patterns which are exclusively used in Chinese words.

Stage three: Positioning of stroke-patterns

At a more advanced stage, children proceed from the gross features of the words to finer features in writing. Five-year-olds have begun to receive formal instruction in reading and writing and they have learned some strokes and stroke-patterns. Lai (5 yrs 3 mth) could write her name and the word *mother* perfectly. She did not know how to write *house* and *elephant*, but she managed to write the first character for the word *elephant* and the first stroke for the word *house* correctly. The sequence of strokes is very important in the Chinese writing process. She had also positioned all the stroke-patterns in squared shapes, and all the words, disregarding the complexity of stroke-patterns, were of similar size.

In this study, the words written by the children were measured by their superordinate and ordinate graphic features. Superordinate graphic features refer to linearity, presence of units, regularity of blanks and constricted size. Each production was also classified according to its ordinate graphic features. The categories for ordinate graphic features were (1) unidentifiable scribbles; (2) pictorial representation; (3) pictorial but squared in shape with constricted size; (4) approximate writing and (5) correct writing. A significant age difference was found in all the writing tasks. Name writing was more advanced than writing other unfamiliar words - *house* 〈屋〉, *mother* 〈媽媽〉, and *elephant* 〈大笨象〉 in all age groups. It is noted that although the 3-year-olds use drawing as their basic writing model, their name writing was at a more advanced stage than other writing (Chan and Louie, 1992). The 4-year-olds can recognize the distinctive orthographic features in their own names and are able to produce familiar strokes and stroke-patterns. Some of the five-year-olds can make approximate and even correct writing for their own names. The writing of the words, *house* 〈屋〉, *mother* 〈媽媽〉, and *elephant* 〈大笨象〉 appeared to be too difficult for most of the children. More clues are needed to lead children to use different strategies in writing unfamiliar words.

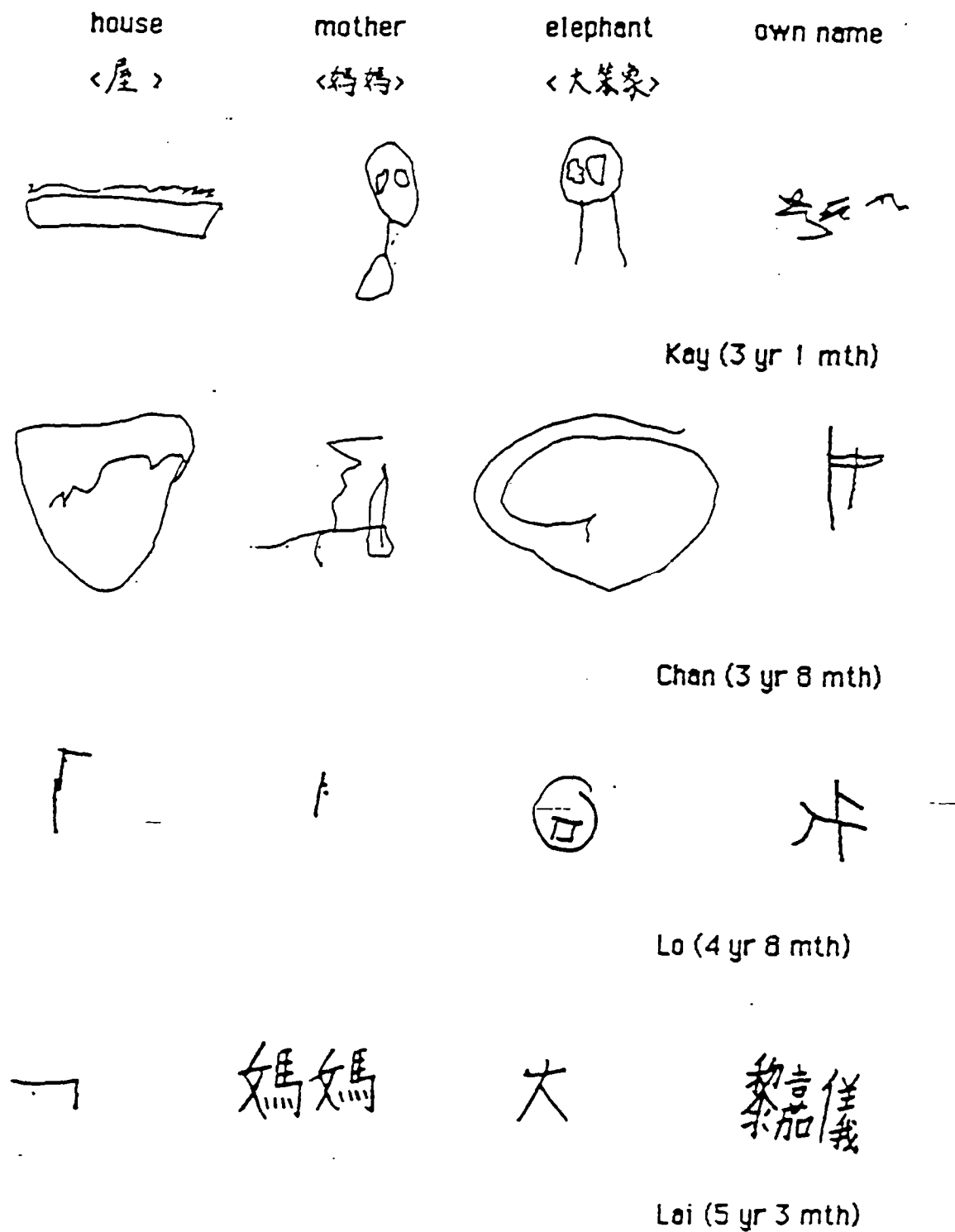


Figure 4.1 Examples of preschool children's writing

The study suggests that young children make active attempts to understand the written language. At age four, they have differentiated picture and print. They apply strokes and stroke-patterns in writing which are the basic units of Chinese writing. What is still unclear is the kind of strategies children will develop when they are formally taught to learn. With the exposure to many new words, children need other strategies to understand the written system in order to help them learn efficiently. This study failed to provide any evidence on this aspect as most of the older children refused to write anything apart from words that they had been formally taught.

4.3.1.2 Developing global visual skills in reading Chinese

Another reason why children progress in their understanding of written Chinese is attributed to the development of visual skills which enables them to make finer distinctions between written words. There is a large number of visually distinctive, complicated words in Chinese. It is likely that some visual discrimination and visual memory skills are important for learning to read Chinese.

Ho (1994) was interested to find out the kind of visual skills needed for learning to read Chinese. The main objective of her study was to discover the precursors of learning to read Chinese; in particular, whether visual skills are significant precursors of learning to read. One hundred Chinese children from the age of 3 years 4 months, participated in the two-and-a-half-year longitudinal study. Predictive measures were taken before children could read any Chinese. Children were tested extensively on various visual skills. One standardized test (The Frostig Developmental Test of Visual Perception) and two specially-constructed tasks (a Visual Detection of Matching Parts tasks and a Visual Matching task) were used to assess the children's visual discrimination skills. A Visual Sequential Memory task and a Memory for Abstract Designs task were constructed as measures of visual memory skills. The Frostig Developmental Test of Visual Perception is a test of various visual perceptual abilities and there are five subtests - Subtest I (Eye-motor coordination), Subtest II

(Figure-Ground), Subtest III (Constancy of Shape), Subtest IV (Position in Space) and Subtest V (Spatial relationships). For the outcome measures, which were conducted during the first two years of the children's reading instruction, a reading and a writing task were constructed to assess children's literacy skills in Chinese.

Results indicated that the scores of Frostig Subtest III (constancy of shape) were the most significant and consistent predictor of children reading Chinese words at age four and five years. The visual memory skills measured by the Memory for Abstract Designs task were also significant predictors of the children's word reading at the age of five. As the predictive visual measures were taken before the children could read any Chinese, and the effect of important extraneous variables (age, IQ, and socio-economic background) were partialled out, the author believed that it is likely, though not definite, that visual discrimination and visual memory are very important causal factors for learning to read Chinese.

It is surprising to find that actually not many different visual skills are needed for reading Chinese. The importance of shape constancy ability for reading Chinese is probably due to the fact that children have to detect recurring stroke-patterns, which are the basic constituents of Chinese words. Visual memory skills are important for learning to read Chinese probably because children have to remember a large number of visually complicated and similar words. In spite of this, visual memory skills are important in reading for a relatively short period of time. Out of the three measures which were taken at 4;4 years, 5;3 years and 5;8 years, visual memory skills were a significant predictor of Chinese Word Reading at the second time (5;3 years) only, not on the first (4;4 years) and the third occasions (5;8 years). More evidence is needed to confirm the general belief that visual skills are particularly important in learning to read Chinese.

4.3.2 Phonological aspects of written Chinese

4.3.2.1 Developing one-to-one correspondence between a syllable and a character

Chinese is a syllabic language, with one syllable corresponding to one character. Chan (1990) set up a reading task to examine if preschool children utilized syllabic clues to help them to read. Sixty preschool children, aged 2 - 6 were invited to take part in a matching task. The task consisted of 18 sets of pictures with four corresponding word cards. A familiar object, e.g. a table or a camera, was printed on each picture card. Out of eighteen objects, there were six objects with names of one-character-words, six with names of two-character-words and six with three-character-words. In each set, there were four word cards to match the picture. The four word cards contained one correct match and three wrong matches and children were to guess which card had the name of the object in it. Since one character corresponds to one syllable in every word, it was predicted children would use the syllabic clue to help them choose the right card as they would not be able to read most of the words on the cards.

Results from the matching tasks showed that the three-year-olds had not yet acquired the syllabic rule. They paid little attention to the word cards and had a tendency to pick up the card nearest to the picture as the name of the object. Four-year-olds spent more time reading the words but only one child applied the syllabic rule in choosing the word cards. Others chose the cards at random; some of them picked up only one-character words, disregarding the name of the objects. It appears that they were trying to form some hypotheses but no consistent patterns had yet been formed. Many five-year-olds enjoyed the matching task. When they were required to choose a word card to match the picture, they adhered to the syllabic hypothesis in choosing the correct one. Some children counted the number of characters on each word card before they decided on the correct response, which indicated clearly that they had already grasped the one-to-one correspondence between a syllable and a character.

Lee (1989) also had similar findings , that is, at the age of five, children can apply the syllabic rule in reading environmental print. In her study, most of the 5 and 6 year olds and even some of the 4 year olds demonstrated in their responses that they had the concept of one-to-one syllable-character correspondence for the print presented in the study.

In the following example, the child, Qi, has applied one-to-one correspondence in his reading:

TEXT REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



(Lee 1989, p.209)

From the above example, we can see the effort children make trying to regulate the system step by step. Qi (6 yrs old) has acquired the one-to-one correspondence rule,

and at the same time, he has acquired by himself, another rule which is not appropriate in Chinese, that is, words with the same sound should be written in the same way. Although all the Chinese characters are one syllable long, a subtle change in the consonant and vowel and even the tone will lead to another word. In addition, there are a lot of homophones in Chinese which means that the same sound can be represented by different characters. Many young children find it difficult to understand why the same sound can be represented by different words. Yang (1990) found that the most frequent errors children made in their spontaneous writing was the substitution of a word with the same or similar pronunciation which accounted for 60% of the total errors.

4.3.2.2 Developing global phonological skills in reading Chinese

Ho (1994) investigated whether good phonological skills have positive effects on learning to read Chinese, as differentiating Chinese words demands acute attention to the variation in sound. In her two-and-a-half year longitudinal study, 100 Chinese children from age 3;4. were given two predictive phonological measures and outcome measures were taken on three occasions to examine the causal link between phonological and literacy skills.

Predictive phonological measures included two phonological awareness tasks and a phonological memory task. The two phonological awareness tasks include a Partial Homophone Detection task (e.g. [sau]2, [tin]1, [tin]4); and a Rhyme + Tone Detection task (e.g. [fa]1, [mun]4, [tsa]1). Children were asked to choose the odd sound among the three spoken syllables presented together with pictures in all the three tasks. For the phonological memory task, children were required to repeat syllable-strings (rhyming and non-rhyming) spoken by the experimenter. For the outcome measures, several types of tasks were constructed to assess the children's literacy skills in Chinese, namely (1) Chinese reading; (2) Chinese writing; and (3) Chinese Pseudoword Reading. Single-character word reading, two-character word

reading and reading comprehension were constructed for the Chinese reading tasks. Children were also asked to write single-character words and two-character-words. Invented words were used in the Chinese pseudoword reading task.

The results showed that phonological awareness was a significant predictor of Chinese word reading at 5;3, but not at 4;4 years or 5;8 years. The Chinese Partial Homophone Detection and Chinese Rhyme and Tone Detection were significant predictors of Chinese Word Reading at 5;3.

4.3.2.3 Applying Orthographic-Phonology Correspondence (OPC) rules in reading Chinese

Over 90% of Chinese characters consist of two components - the semantic radical and the phonological component. Pronunciation can be achieved by analogy or by derivation via the phonological component. There are signs that preschool children can utilize the phonological components for pronunciation. A Chinese Pseudoword Reading Task was conducted by Ho (1994) to assess the ability of five-year-old children in reading new Chinese words by applying their knowledge of OPC rules in Chinese. As the pseudowords have no lexical identities, the best possible way of reading them is by using OPC rules. Twelve pseudowords were made up and children were asked to guess the pronunciation. Each pseudoword was made up of a radical and a phonological component in a left-right structure and in their legal positions. If the children knew that the phonological components could provide phonological clues to the pronunciation of the word, they should read the Chinese pseudowords according to the pronunciation of their phonological components (by derivation), or according to the pronunciation of other words sharing the same phonological components (by analogy).

Percentages of correct responses in reading the pseudowords were 79% for right-side phonological components and only 49% for left-side phonological components, suggesting that the five year old children may have acquired some rudimentary knowledge of the OPC rule. The better performance in reading pseudowords with right-side phonological components than with left-side phonological components was probably related to the fact that the majority of Chinese words have phonological components on the right side. This positional regularity of phonological and radical components helps children to realize that the right-side components are usually related to the sound of a word.

No phonological awareness task was a significant predictor of Chinese pseudoword reading, and in return, the Chinese pseudoword reading was not a significant predictor of either Chinese Rhyme Detection or Chinese Tone Detection which were conducted in the later outcome measure sessions. Furthermore, Chinese pseudoword reading did not correlate significantly with any other reading tasks after controlling for the effect of differences in IQ. The author suggests that the five year old children may not apply their rudimentary knowledge of the orthographic-phonology correspondence (OPC) rules actively in reading real Chinese words.

The mean score of Chinese pseudoword reading was 7.51 out of 12 items, which means that five year old children can read at least half of the 'new' words by themselves. One of the reasons why Chinese pseudoword reading did not correlate with the Reading tasks is because in the reading tasks, children are required to read one-character-words, two-character-words and short sentences, and many of the one-character-words are simple characters where OPC rules are not applicable. The OPC rules are only applicable for compound words with semantic radicals and phonological components. Therefore, a more relevant reading task is needed in order to examine how children apply their knowledge of OPC rules to reading.

It appeared that, apart from global phonological awareness, children at the age of five begin to develop another skill which allows them to look into the components of a word for phonological clues. Over 90% of Chinese characters consist of a non-radical component which may provide a phonological clue to the whole word. This knowledge, although not explicitly taught in school, seems to be acquired by children in the process of learning to read.

4.3.3 Semantic aspects of written Chinese

4.3.3.1 Understanding that print conveys meaning

One of the essential skills children have to acquire in reading is to get the meaning from print. Lee (1989) examined how Chinese preschool children read and responded to environmental print with different degrees of decontextualization. She predicted that there would be a developmental growth in the understanding of the function of print even before children had been formally introduced to reading. Sixteen kindergarten children, aged three to six were invited to read 20 familiar environmental print items in contextualised and decontextualised situations. She was interested to know if preschool Chinese children showed an awareness of the print in their environment. She utilized Print Awareness Tasks devised by Goodman (Goodman, 1984) with the adaptations necessary to the Taiwanese setting and the Chinese language. Twenty items which included print from the environment familiar to young children in Taiwan were chosen. They included labels from household food containers or packages, street signs, signs of popular stores, toy packages, TV commercials, or public services which are commonly seen in daily life. In the first session, children were invited to read the print in full environmental context. For example, a cut-out portion of a container or a picture of a street sign was shown together with the words. In the second session, children were asked to read the print again without the support of the context.

Results indicated that about 30% of all of the three-year-olds can read environmental print when it is embedded in context. By age six, 60% of the subjects can get it right. Fewer subjects are able to read the print at any point of decontextualization (only 16% for the three-year-olds and 54% for the six-year-olds). Subjects also use their hypotheses about the nature of the Chinese writing system to read the print items. These hypotheses include syllabic correspondence, character recognition, and graphic feature analysis. Most of the subjects demonstrate the use of the syllabic hypothesis in their responses. To a lesser degree, the children also use character recognition which appears to be the last strategy which subjects choose during the process of responding. A few of the subjects use graphic feature analysis to identify the items, and two subjects showed evidence of referring to the semantic radicals in reading the print.

There were cases where the subjects could already focus on character identification, although they could not yet give a meaningful answer.

TEXT REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



(Lee 1989,p.169)

Han (6 yrs old) could recognize a few characters but she failed to read the print because conflicting information provided by the context could not be accommodated. There are also indications that older children, in order to get to the meaning, paid attention to graphic details that were smaller than characters. The study did not

demonstrate whether young children are aware that the semantic radicals convey any meaning in understanding a character, a word or a phrase.

4.3.4 Conclusions

To conclude, these studies have demonstrated that although preschool children aged three to five have not yet been formally taught to read, they already have some knowledge about the writing system. Their writing has evolved from drawing to applying conventional strokes and stroke-patterns. Seventy-five percent of the five year old children in Chan's (1990) study had discovered the one-to-one correspondence rule between a character and a syllable in reading. Ho's (1994) study has provided evidence that five-year-olds can even apply rudimentary OPC rules in reading some compound words. In searching for the meaning of the words, they gradually pay attention to characters and semantic components. All these findings support the idea that children are active in learning to read and write, and that they construct rules to interpret the written system. We expect similar active involvement when children are formally introduced to reading and continuously develop more rules that can help them to understand the system.

At the same time, we can trace developmental growth even amongst preschool children in their understanding of the orthographic, phonological and semantic aspects of written Chinese. Older children perform better than younger children in the sense that they have developed a more conventional way of reading and writing. In the present study, I propose that older children will develop reading strategies, which are developed incidentally by adult readers, when they are formally introduced to reading. Research looking at the skills of fluent readers has demonstrated that adults read analytically, rather than holistically, in getting to the meaning and sound of words. In the following section, research with older children who have been formally introduced to reading is reviewed. Most of the research has adopted a holistic approach, regarding a character as the smallest unit in reading. Only one piece of research (Ho,

1994) has studied how children utilize the phonological components and semantic radicals in reading.

4.4 Literacy development at the formal reading stage

4.4.1 Orthographic aspects of written Chinese

4.4.1.1 Developing global visual skills in reading Chinese

Lee, Stigler, and Stevenson (1986) attempted to investigate the rates at which children acquire skills in reading English and Chinese, and the kinds of skills related to early progress in reading different orthographies. As part of their study, 912 Chinese six-year-old first graders in Taiwan were given an IQ test, three reading tests, a cognitive test, and two visual tests. The two visual tests were a coding task, which involved detection of spatial differences of up-down and left-right relations; and a spatial relations test, adapted from the Thurstone Primary Mental Abilities Battery which emphasized detection of differences of small details. It was found that the scores on the spatial relations test correlated significantly with the Chinese composite reading scores. The study failed to provide a strong and specific relationship between visual skills and reading because only simple correlations were found. IQ was not partialled out, and it was unclear whether the relationship would still hold after controlling for this effect. Also as the different reading scores (word reading, text reading, and reading comprehension) were combined into a composite score for analysis, it is not clear whether spatial relations are particularly related to word recoding, or to reading comprehension, or to both.

In another cross-cultural study, Huang and Hanley (1992) investigated the relationship between visual skills and reading performance for English children in Britain and Chinese children in Hong Kong and Taiwan. One of the aims of the study was to analyze the relationship between visual skills and reading in English

and Chinese and to investigate whether or not any of these factors can predict the children's success in reading English and Chinese. For the Chinese subjects, there were altogether three groups who were all eight-year-old Chinese third-graders from Hong Kong and Taiwan. Sixty subjects from Hong Kong were equally divided into the Pin-Yin group (those who have learned Pin-Yin for about one year) and non Pin-Yin group (those who have not learned Pin-Yin and cannot speak Mandarin). In addition, there was a Taiwan group which consisted of 50 children from Taiwan who had learned Pin-Yin and spoke Mandarin. They were given a Chinese word reading test, an IQ test, a verbal vocabulary test, two visual tests, and two phonological tests. The two visual tests included a visual form discrimination task and a visual-paired associates task.

The results from the two visual tests were inconsistent. The visual form discrimination task correlated significantly with reading in the Taiwan group, but not with any of the Hong Kong groups; the visual-paired associates task correlated significantly with the Taiwan group and the non Pin-Yin group in Hong Kong, but not with the Pin-Yin group in Hong Kong. With respect to predicting Chinese reading abilities, visual skills predict Chinese reading ability for the non Pin-Yin group and Taiwan group but not the Pin-Yin group. It is difficult to explain the discrepancy in the results as all the subjects in the study had learned complicated characters. Being able to speak Mandarin or the knowledge of Pin-Yin should not affect visual skills.

Huang and Hanley (1993) have also conducted a one-year longitudinal study in Taiwan attempting to examine the relationship between phonological awareness and visual skills before children received formal reading instruction, and subsequent reading ability in Chinese in the first school year. One of the hypotheses was that there would be a significant correlation between early visual skills when assessed before children received formal instruction and subsequent Chinese reading development. Forty six-year-old children were tested three times in their

first grade: the first time before learning Pin-Yin and Chinese words (Time 1), the second time after learning Pin-Yin but before learning Chinese words (Time 2), and the third time after learning Pin-Yin and Chinese words for one school year (Time 3).

At time 1, the children were given an IQ test (Raven's Coloured Progressive Matrices), a Chinese word reading test, two phonological tests and a visual test. They were retested with the same battery of tests at Time 2 and at Time 3. A visual-paired associates test was given to test children's visual skills three times during their first school year. In the visual-paired associates test, children first learned the colour associated with each of six abstract line drawings. They were then presented with each of the line drawings in turn and asked to indicate which colour went with each figure.

It was found that the simultaneous correlations between scores on the visual-paired associates test and those on the Chinese word reading test at all three testing points were not significant after adjustment for the effect of IQ. In addition, the visual scores failed to predict reading scores ten months later, after the effects of IQ and phonological test scores were partialled out. This result was unexpected to the authors, as Huang and Hanley (1992) had found that reading ability in third grade Taiwanese children was significantly correlated with visual skills, measured by the same visual-paired associates test.

Huang & Hanley (1993) failed to demonstrate the importance of visual skills in beginning reading, and Ho (1989) also failed to find a reliable connection between visual skills and early reading. She gave Chinese second-graders in Hong Kong two visual tests (spatial relations and memory for abstract designs). Her hypothesis was that visual skills were correlated with Chinese reading. In each trial on the spatial

relations test, two identical abstract designs and one visually similar design were shown. Children were required to pick out the two identical abstract designs. In the memory for abstract designs test, subjects were asked to recognize the target from six options presented immediately after the target designs.

It was found that neither visual test score was significantly correlated with Chinese word reading or with sentence comprehension after the effects of differences in IQ were partialled out. Ho concluded that the visual skills that she measured did not make any independent contribution to reading for the Chinese second-graders once the contribution of IQ had been controlled.

Chen, Lau and Yung (1993) conducted a large-scale cross-sectional study in Hong Kong to examine the significance of different component skills in reading Chinese in the primary school years. Different component skills - physical feature processing skills (visual and phonological), syntactic knowledge and semantic analysis - were measured at different grades. They expected that different component skills might be more dominant at different educational levels. A total of 622 Chinese children from first to sixth grade were given a Chinese reading proficiency test, a visual analysis test, a phonological analysis task, and some syntax knowledge and semantic analysis tasks. The Chinese reading proficiency test consisted of five parts: utilization of words, use of punctuation, organization of sentences, passages for comprehension, and composition.

A visual analysis task was designed to tap children's skill in using Chinese orthographic regularity. Research with English readers (e.g. Massaro, Venezky and Taylor, 1979) has found that sensitivity to orthographic regularity is well developed among good readers, and it was predicted that sensitivity to orthographic regularity is also important in reading Chinese. Children were required to detect a

designated component of a stroke-pattern (□ mouth) in four types of passage; a passage of prose, a passage with scrambled words, a passage with pseudowords and a passage with nonwords. The stroke-pattern can be a word by itself or it may be the component of a character. It was predicted that the degree of difficulty in detecting the designated stroke-pattern would increase if the passage is more meaningful. The different passages were used to increase diversity of performance. Texts for the passage with scrambled words were first chosen from normal texts and then the order of words in each sentence was scrambled so that the word string did not convey any meaning. The passage with pseudowords consisted of made-up words. These were not real words but they looked like real words because the stroke-patterns were in their legal positions. The passage with nonwords consisted of items which were neither real words, nor did they look like real words. The nonwords were made by violating the orthographic rules, such as the violation of legal positioning of stroke-patterns. These four types of passage were presented in blocks in a booklet form and in each case, subjects were asked to detect the component (□ mouth). All children were given seven minutes to complete the task. The number of targets in each passage varied from 14 to 17 in grade one and 22 to 29 in grades two to six.

Results from the study showed that children generally performed very well on the visual analysis task. The first and the second grade children, in particular, did best on visual analysis as compared with their performance on different component skills, and they even scored higher than the third, fourth and fifth grade children on this task. This suggests that sensitivity to orthographic regularity may be particularly important for young children when they have to learn many new words. The skill becomes relatively less important in the upper grades as there are no developmental differences in visual analysis across all grades, whereas in other skills such as phonological analysis, syntax knowledge and semantic analysis, older children always performed better than the younger ones. The authors also found that subjects were more likely to miss the target component when the passage was

more meaningful and when the test items were more word-like. This implies that orthographic regularity is helpful in decoding unknown rather than known words.

In a series of simultaneous multiple regression analyses, it was found that the relationship between visual analysis and reading was not significant at any grade level, which is not surprising because the study measured various aspects of Chinese reading proficiency such as use of punctuation, organization of sentences, passages for comprehension, and composition. Visual analysis is more likely to be correlated only with decoding words.

To conclude, global visual skills have been measured by various tasks: a spatial relations task and spatial difference task (Lee, Stigler, & Stevenson, 1986); a visual form discrimination task and visual paired-associate task (Huang & Hanley, 1992); and a visual analysis task (Chen, Lau and Yung, 1993). The results from these studies failed to provide evidence that visual skills, in general, are important for reading Chinese. It is possible that global visual skills are not particularly needed in reading familiar words, but it is difficult to believe that visual skills are not needed in decoding new words. Although the visual processing skill was found to be irrelevant to reading proficiency in the study, it is still not known whether visual analysis of stroke-patterns can help children to decode words efficiently.

There is reason to believe that visual analysis of stroke-patterns, rather than global visual skills, enables children to understand the orthographic structure of written Chinese, which in turn helps them to learn to read. Research on skilled readers has indicated that Chinese adult readers use stroke-patterns in the visual analysis of words (Chen, 1993). Ho (1994) found that constancy of shape (one of the visual discrimination skills) was the only significant predictor of reading Chinese at four and five years, after controlling for the effects of differences in age, IQ and mothers'

education. Constancy of shape was designed to test the ability to detect particular shapes (circles and squares) embedded in and mixed with other visually distracting figures and the author suggested that an equivalent of shapes in a word would be stroke-patterns which serve as the constituents of individual Chinese words. The same stroke-pattern, though recurring, could appear in different sizes at different locations of different words. Chinese word reading requires the identification of the stroke-patterns appearing in different words, such identification being similar to the constancy of shape task. Therefore, the ability to detect constancy of shape as an important predictor of beginning reading success in Chinese seems to reflect the importance of detection of stroke-patterns in Chinese word recognition.

4.4.2 Phonological aspects of written Chinese

4.4.2.1 Developing global phonological skills in reading Chinese

So and Siegel (1992) conducted a cross-sectional study to examine the relationship between phonological, semantic and syntactic skills and reading ability in Chinese. Focusing on grades 1 to 4, and using Chinese children whose mother language was Cantonese, the aims of the study were (1) to identify the relationships between reading ability, phonological, semantic and syntactic skills and short-term memory; (2) to uncover developmental patterns in those skills; (3) to compare normal and poor readers on these skills. One hundred and ninety-six children in Hong Kong were tested on a Chinese word reading test, two phonological tests (tone discrimination and rhyme discrimination), and some semantic and syntactic tests. The children who scored in the lowest quartile of the Chinese word reading test were classified as poor readers. In the tone discrimination task, four Chinese homophonic syllables, two of the same tone and two of different tones (e.g. [hon]³, [hon]², [hon]⁴, and [hon]³), were orally presented to the children. The children were asked which syllables had the same tone. In the rhyme discrimination task, four Chinese syllables, two of which rhymed with each other (e.g. [fuk]¹, [huk]¹, [fong]⁴,

and [bei]⁶), were presented orally. The children were asked which two syllables rhymed with each other. There were 15 trials altogether in each task.

It was found that the normal readers at each grade performed significantly better than the poor readers in the tone discrimination task and the rhyme discrimination task. However for normal readers, there were no significant differences between grades on tone and rhyming tasks dealing with phonological skills. When the effect of age differences was partialled out, the two phonological tasks correlated significantly with reading scores. Although there is evidence that poor readers at each grade performed significantly lower than normally achieving readers, the study failed to show any developmental patterns in phonological awareness, as there is no significant difference between grades on tone and rhyming tasks dealing with phonological skills for normal readers. The authors concluded that for the normal readers, phonological skills may be equivalent to phonemic awareness which is acquired in the initial stage of reading development. However, out of the 15 trials, children from Grade 1 to Grade 4 got approximately two thirds of the responses correct; the lack of age difference was not due to a ceiling effect. If sensitivity to tone and rhyme are crucial, why did older children fail to advance further in these two aspects?

In another extensive study tracing the development of children's reading in Hong Kong, one of the aims of Chen, Lau and Yung (1993) was to examine whether the ability to use a phonological code to access the lexicon is strongly related to early reading acquisition. In addition, they were interested to find out if phonological analysis skills were important in all grades in the primary school years. Six hundred and twenty-two children from first to sixth grades were given a phonological analysis task, a Chinese Reading Proficiency test, and some other tests to examine other component skills. The phonological task was designed to test children's ability to transcribe sounds from meaningless homophones into words. The task was analogous to the "brane-brain" type of task where subjects who saw "brane"

were asked to write down a proper word "brain". In this task, the subjects were presented with a pair of characters which form a meaningless homophone to a real Chinese two-character word, and were asked to write down as quickly as possible the corresponding real compound word.

Results from the study showed that phonological analysis was a significant predictor of children's reading performance at grades one to three, but it did not make any substantial contribution to the prediction of reading performance at higher grade levels. The presence of homophones is particularly difficult for young children as they cannot understand why the same sound can be written in different ways.

In another study, Ho (1989) attempted to explore the relationship between reading, phonological and visual skills for Chinese beginning readers. The rationale of her study was that there was a relationship between phonological and visual skills and the reading ability of second-graders. Forty-seven eight-year-old Chinese second-graders were given an IQ test (Raven's Standard Progressive Matrices), a Chinese word reading test, a Chinese sentence comprehension test, a rhyme detection test and two visual tests. In the rhyme detection test, three Chinese syllables, two of which rhymed with each other (e.g. [so]¹, [bo]¹, and [hung]⁴), were presented orally to subjects. The children were asked which two syllables sounded similar. It was found that rhyme detection correlated significantly with both Chinese word reading and sentence comprehension after the effect of differences in IQ scores was controlled for. The author concluded that rhyme awareness is important in learning to read Chinese.

Research on phonological awareness in Chinese children has demonstrated that there is a correlation between phonological awareness and reading ability. In

Chinese, there is a one-to-one correspondence between a character and a syllable, and a subtle change in the consonant, the vowel and even the tone of the character will lead to another word. Tone and rhyme discrimination are needed to make acute differentiation in the sound of the words. Ho (1994) also suggested that phonological awareness can help to make connections between the sounds of the whole word and its components, as many words share the same/similar pronunciation with their phonological components.

4.4.2.2 Applying Orthographic-Phonology Correspondence (OPC) rules in reading Chinese

Ho (1994) investigated the importance of the knowledge of the orthographic-phonology correspondence (OPC) rules in reading for Chinese beginning readers. Chen's (1993) study demonstrated that Chinese adult readers rely on the phonological components for pronunciation. In a study on preschool children, Ho (1994) found that five year old children have already acquired some rudimentary knowledge of the OPC rules in Chinese although they cannot apply the rules actively in reading. Assuming that more words have to be learned in order to be able to generate the OPC rules, Ho examined the utilization of phonological components for pronunciation amongst the first and second graders. Children were asked to read Chinese ideophonic compound words and pseudowords. If the OPC rules have psychological reality for the children, then congruent (i.e. phonologically regular) Chinese ideophonic compound words should be read better than incongruent ones, especially when the words are infrequent. Furthermore, as no lexical information is provided by the pseudowords, reading Chinese pseudowords should largely depend on the children's knowledge of the OPC rules. In addition, it was also predicted that phonological awareness such as onset and rhyme detection are important for the acquisition of the OPC rules. A Chinese word reading task was administered to assess the children's reading skills.

Forty-five grade one and forty-five grade two children took part in this study. The following tasks were conducted: 1. Raven's Standard Progressive Matrices; 2. Chinese word reading (50 Chinese two-character words from grade one to grade four level); 3. Chinese ideophonetic compound word reading (60 ideophonetic compound words of two level of frequency and three levels of congruency) 4. Chinese Pseudoword Reading (17 pseudo-ideophonetic compound words) 5. Chinese radical detection task 6. Phonological awareness tasks (Chinese onset detection and Chinese rhyme detection).

The results showed that the Chinese Rhyme Detection task correlated significantly with the Chinese Ideophonetic Compound Word Reading task and the Chinese Pseudoword Reading task for the first graders, but not for the second graders. The author suggested that for the first graders, rhyme awareness may be important for acquisition of knowledge of the OPC rules. However, for the second graders, once knowledge of the OPC rules has been acquired, rhyme awareness may no longer be important for the application of these rules in reading.

The results of the Chinese Ideophonetic Compound Word Reading task showed that there were higher rates of correct responses in reading phonologically congruent than incongruent Chinese words. The significant congruency effect in reading Chinese ideophonetic compound words suggests that first and second grade children can make use of the phonological component of a word to facilitate reading. Error analyses show that "phonetic derivations" and "phonetic analogies" were the two most common types of error. They accounted for 62% of all errors. This therefore confirms the psychological reality of the OPC rules for Chinese beginning readers.

The Chinese Word Reading task and the Chinese Ideophonetic Compound Word Reading task correlated significantly with each other in Grade 1 and Grade 2 after controlling for the effect of differences in IQ. The Chinese Pseudoword Reading task also correlated significantly with these two reading tasks at both grade levels. Thus, knowledge of the OPC rules was important for both grades in reading Chinese ideophonetic compound words and for their reading of Chinese two-character words.

There are however a few problems with testing young children's knowledge of OPC rules by reading ideophonetic compound words. The orthographic structure of the compound words has been highlighted because two-thirds of the stimuli are left-right ideophonetic compound words and the phonological components can always be found on the right side of the words. It is difficult to know whether the children are referring to rules of positional regularity of phonological components (that is, the phonological component is always on the right for left-right ideophonetic compound words, and at the bottom for top-bottom ideophonetic compound words), or if they can really utilize the non-radical component to pronounce the words.

The results from this study suggest that phonological awareness, apart from facilitating children's differentiation of words with similar pronunciation, also helps them to read words analytically, by looking into the phonological components for pronunciation. This kind of analytical skill is believed to be particularly important in the process of learning to read, as children can get to the sound of the word by themselves by following the OPC rules.

4.4.3 Semantic aspects of written Chinese

4.4.3.1 Semantic understanding of words

In examining the relationships among reading ability, phonological, semantic and syntactic skills in Chinese, one of the aims in So and Siegel's (1992) study was to identify the relationship between reading ability and semantic understanding of words and to trace the developmental patterns of semantic processing among Grade 1 to Grade 4 Chinese children in Hong Kong. One hundred and ninety-six children were given a reading task and a word meaning task. In the word meaning task, children were required to choose a two-character-word out of three alternatives which had similar meanings to the target word. Fifteen experimental trials were administered. Results from the study showed that there was a high correlation between the word meaning task and the reading task indicating that semantic understanding of words is crucial in reading Chinese. A developmental trend in semantic processing was found, the mean scores in Grade 3 and 4 were significantly higher than Grade 1 and 2 and Grade 2 was higher than Grade 1. In this study, the target words and the chosen words were read to the subjects orally without showing them the print, and thus it is unclear if the semantic radicals contribute to the semantic understanding of the words.

4.4.3.2 Detecting semantic radicals

A Chinese radical detection task was conducted by Ho (1994) to measure children's orthographic knowledge of semantic radicals. It was predicted that detecting the radical component of a word may be necessary for orthographic decomposition and application of the OPC rules in reading Chinese, as radical detection indirectly helps identification of the phonological components.

Twenty Chinese words of first grade level were selected. Half of the semantic radicals in the selected words were simple words on their own, and half of them were just stroke-patterns. The position of the radicals could be found at the top, at the bottom, on the left or on the right side of the target words. The 20 words were printed in two columns on a piece of paper, and children were asked to look for the semantic radicals in each word. Results showed that both Grade 1 and Grade 2 children could get 14 words out of 20 correct. However, Chinese radical detection did not correlate significantly with the Chinese pseudoword reading task or with Chinese word reading for the seven- and eight-year-olds. The author suggested that it is possible that orthographic decomposition in Chinese word reading does not depend on radical detection (to identify the non-radical part as the phonological component), but depends directly on identifying the phonological component itself. Thus, the result concludes tentatively that the ability to detect the radical component in a word does not seem to be important for the acquisition of knowledge of the OPC rules in Chinese. A replication study is needed to explain why the detection of semantic radicals, as part of the process of lexical decomposition, takes place with adult readers (Chen 1993) but not with young children.

4.5 Conclusion

4.5.1 Orthographic aspects of written Chinese

The findings from the research enable us to draw a general conclusion as to how children come to recognize and remember words. Initially children distinguish the basic strokes and stroke-patterns exclusive for written Chinese (Chan 1993) and then they pay attention to the positions and recurring nature of stroke-patterns (Ho, 1994). Visual memory was found to be important for learning to read but only for a short period of time (Ho, 1994) as children soon found out that very few words can actually be remembered purely by visual memory. This has also been shown in research on Chinese fluent readers (Zhang and Simon, 1985; Hue and Erickson,

1988). Children have to devise other ways to help them identify large numbers of new words which they encounter daily.

There are very few traces of a relationship between the development of global visual skills and reading which implies that global visual skills may not be of great importance in reading Chinese. Instead, I propose that understanding the orthographic structure of a word might be the crucial skill to help children to recognize and remember words quickly and efficiently. Chen (1993) has provided evidence that adult readers use stroke-patterns as the unit of visual analysis and Ho (1994) has suggested that being able to identify the recurring nature of stroke-patterns is important in early reading. If children pay attention to the stroke-patterns, they can at some point realize that stroke-patterns are the building blocks in over 90% of the Chinese characters. The graphic structure of the stroke-patterns has been standardized and many of them occupy fixed positions within a compound character. If a developmental trend can be traced in children's understanding of the graphic structure and positional regularity of stroke-patterns, it will provide strong evidence that children are actually using the stroke-patterns as units of visual analysis in helping them to learn to read Chinese. No study, however, has ever looked into the importance of stroke-patterns in the way children learn to read.

4.5.2 Phonological aspects of written Chinese

There is a one-to-one correspondence between a character and a syllable in Chinese, and six year old children have already acquired this knowledge when they begin to learn to read (Chan, 1990; Lee, 1989). However, the presence of homophones, which is a dominant feature in Chinese, poses difficulty as children cannot understand why the same sound can be represented by different words with different graphic features and different meanings. In addition, children have difficulty in paying attention to the subtle differences in pronunciation (Yang 1990), as a slight difference in consonant, vowel and even in tone would result in a

different Chinese word. Global phonological skill is correlated to reading (Ho, 1989). Developing global phonological skill helps children attend to minor differences in sound so that they can differentiate words more efficiently. Chen, Lau and Yung (1993) found that upper graders acquire better global phonological skills than Grade 1 to Grade 3 children. No significant difference in the development of phonological skills was found in lower grades (So and Siegel, 1992).

On the other hand, Ho (1994) has observed significant differences in the acquisition of the OPC rules between first graders and second graders. The Orthographic-Phonology-Correspondence (OPC) rules enable the pronunciation of a word to be reached via the phonological component. The rules include a lexical decomposition rule - dividing the whole word into a semantic radical and a non-radical component - and locating the phonological component; which, by rule of thumb, is the non-radical component (Chen, 1993). Pronunciation via the phonological component can be reached either through derivation (pronunciation of the whole word is directly derived from the pronunciation of its phonological component) or by analogy (pronunciation of whole word is deduced via analogy with other words sharing the same phonological component). Analyzing the phonological components enables children to read new words by themselves, as demonstrated by Ho (1994), there is evidence that children can read pseudowords by applying the OPC rules. Knowledge of the OPC rules is also significantly correlated with reading Chinese for both first and second graders. However, Ho's study failed to explain whether children undertake a lexical decomposition process to get to the phonological component, or whether they rely only on the positional regularity of the phonological components. Moreover, the study has not investigated systematically if children pronounce words by analogy and also by derivation. Since Ho's (1994) study is the only one that has studied the role of the phonological component in the process of children learning to read, further research in this area is needed to clarify the process leading to pronunciation.

4.5.3 Semantic aspects of written Chinese

Very few studies have looked into the semantic processing of written Chinese. Lee (1989) has demonstrated that children are active in getting meaning from environmental print. Semantic understanding of words, in general, is significantly correlated with reading (So and Siegel, 1993). No research has ever investigated the role of the semantic radical which functions primarily as a semantic component. Ho (1994) believed that the ability to detect the semantic radical is necessary to acquire the OPC rules, but Chinese radical detection failed to correlate significantly with pseudoword reading and word reading for the seven- and eight-year-olds in her study.

To conclude, no research has been carried out so far investigating children learning to read Chinese analytically, referring to the stroke-patterns for visual analysis, the phonological component for pronunciation and the semantic radical for meaning. Although fragmented evidence has been found that children do, in fact, gradually acquire this knowledge, a systematic investigation is needed to examine the significance of stroke-patterns in children's understanding of the orthographic, phonological and semantic aspects of written Chinese.

Chapter 5

Method of Study

5.1 Rationale of the study

In this study, I propose that children do not simply learn to read Chinese compound characters by rote memory. When children are exposed to the written language, they implicitly infer some rules that help them understand the Chinese writing system. These rules may be implicit even to adult readers but they play a role in several reading-related tasks. I want to pursue two basic hypotheses:

- a) Chinese children come to read and write analytically, using the same processes that were identified by Chen (1993) in her work with adults. In her study, Chen has demonstrated that adult readers use stroke-patterns instead of strokes as the smallest functional orthographic units in word recognition. Through the process of lexical decomposition, readers refer to the semantic radicals for meaning and the phonological components for pronunciation. Pronunciation is derived via the phonological components either by analogy or derivation.
- b) Children construct the knowledge to read and write progressively and not in an all or nothing fashion.

Four studies were set up to investigate children's knowledge of the orthographic, semantic and phonological aspects of written Chinese. Different tasks were devised in an attempt to examine how children utilize their knowledge with different task demands. The tasks included visual discrimination tasks, judgement tasks, spelling tasks, decoding tasks, and reading for meaning tasks. It was predicted that children require an understanding of the written Chinese system in order to be able to undertake these tasks satisfactorily.

It was also predicted that with a gradual increase in the knowledge of the orthographic, semantic and phonological aspects of written Chinese, children can establish more specific formal and functional constraints in their reading and writing. For example, children will only accept those graphic patterns that fit the orthographic structure of written Chinese. It is also likely that children will pronounce novel words by following the same set of rules in pronunciation. In other words, a more uniform response from children indicates that they are aware of and are operating in accordance with the formal and functional constraints of written Chinese, and in this way, their understanding of the orthographic, semantic, and phonological aspects of written Chinese are demonstrated.

Study 1 (Visual discrimination of Chinese words) was designed to examine whether children use strokes or stroke patterns as units of visual analysis in word recognition. Children were asked to discriminate pairs of Chinese pseudowords which varied in the number of strokes and the number of stroke-patterns. If strokes are the units of visual processing, then children would make more mistakes in the pairs of words with more strokes than with fewer strokes. On the other hand, if stroke-patterns are the units of visual analysis, then words with more stroke-patterns should be more difficult to discriminate than words with less stroke-patterns.

Study 2 (Orthographic acceptability of Chinese pseudowords and nonwords) was designed to investigate whether children simply commit words to memory or whether their knowledge includes some implicit knowledge of the orthographic structure of the words. Children were asked to decide whether a list of pseudowords and nonwords could be acceptable as 'words'. If children remember words purely by rote memory without analysing the orthographic structure of Chinese, then any 'character-like' graphic units should be acceptable as words. If children are analytical in learning Chinese, then they should be able to detect a stroke-pattern in relation to its legal positioning and its legal graphic patterns, and reject words with stroke-patterns which are placed in illegal positions, or written in illegal graphic patterns, even though this knowledge is not taught explicitly in schools.

Study 3 (Creative spelling in Chinese) investigated children's conception of Chinese words in relation to the word constituents - the stroke-patterns. Children were provided with some familiar stroke-patterns and were asked to invent new words as names for some strange objects. If children were aware of the orthographic structure of written Chinese, they would follow the positional rules and use the stroke-patterns to form pseudowords. Some stroke-patterns would be used as the semantic radicals to express appropriate meaning. Other stroke-patterns would be chosen to be the phonological components, and children would refer to them for pronunciation.

Study 4 (Reading Chinese pseudowords) investigated children's ability to utilize the semantic and phonological information in three reading tasks. In the first task, children were asked to choose pseudowords which have a similar meaning or sound to the prototypes, which were real words. If children could locate the relevant linguistic information in different stroke-patterns, then they could transfer this information from real words to read pseudowords. In the second task, children were asked to pronounce some pseudowords. If children understand the functions of the phonological components, then they would refer to them for pronunciation. In the third task, children were asked to choose some pseudowords which would be used as names for some strange objects. If children understand that semantic radicals are used to convey meaning, then they would choose the pseudowords which include the relevant semantic radicals.

5.2 *Setting up the experiments*

All the experiments were specially designed for the present investigation because no study has yet investigated children's understanding of the underlying rules in written Chinese. There were three main characteristics of the experiments: (1) the use of nonwords and pseudowords, instead of real words as clues or stimuli; (2) the adoption of a child-centred approach in designing the tasks; and (3) children's knowledge of the orthographic, semantic and phonological aspects of written Chinese was investigated both in reading and writing.

Nonwords and pseudowords were used in the all experiments. It was assumed that children's underlying knowledge of the written language as a system can best be tested without the interference of known words. Both nonwords and pseudowords are artificially created words. Pseudowords were formed by substituting one stroke-pattern by another in order to create unknown words. Nonwords were created in a similar way but the orthographic structure of the made-up words violated some aspects of word formation in Chinese.

Due to the fact that children were asked to read and write pseudowords and nonwords in the experiments, a meaningful setting for all the tasks was necessary. In the pilot study, various interesting experimental settings were introduced, for example, children were asked to pretend to be teachers; and in another experiment, were invited to take an imaginary journey to another planet. However, although the games were interesting, children could not adjust comfortably to the frequent change of setting of the experiments. In the main study, the presentation of the experiments was modified. All the tasks in the main study were to be conducted in the context of an imaginary situation. With the help of a picture (a boy sitting on a rocket ready to set off) a five year-old boy was introduced. He was going to visit different planets and had to use his knowledge of written Chinese to help the people on the planets to learn to read and write. For example, he had to make orthographic acceptability judgments on some words written by the children on another planet who are learning Chinese. In another

experiment, he was required to create names for some strange objects that had been found on another planet.

Reading and writing demand different cognitive processes. Frith (1985) refers to reading as an input skill and writing as an output skill. Kamberelis (1992) uses the term 'comprehension modality' for reading and 'production modality' for writing. Evidence for the separation of reading and writing strategies in young children is obtained in Bryant and Bradley's (1980) and Kamberelis' (1992) study. Bryant and Bradley (1980) observed that children beginning to learn to read might read words they could not write, and write words they could not read. This odd discrepancy was explained by the fact that the children spontaneously use look-and-say to read words, but used phonic skills to write them. Kamberelis (1992) examined children who are in transition from emergent to conventional literacy and found that there are times in early literacy development when asynchronous relationships between children's knowledge and strategies about reading and their knowledge of and strategies about writing can be found. The present author believes that a feature of the developmental process is that children try to integrate their reading and writing knowledge and strategies into more comprehensive and flexible literacy knowledge and strategies. In this study, reading activities and writing activities were set up to investigate if children's knowledge of the orthographic, semantic, and phonological structure of written Chinese can be utilized in both the input and output processes.

5.2.1 Procedure

A pilot study and a main study were conducted. Both studies were carried out in Hong Kong where Chinese is the mother language for the majority of people living there. There were altogether four experiments in the pilot study and five experiments in the main study. The phonological and semantic constraints in reading Chinese (Study 4) were investigated in one experiment in the pilot study, but it was split into two experiments in the main study. The confounding factors in the pilot study were modified in the main study.

The experiments were conducted in a different sequence across subjects to control for order effects. The experiments in both studies were carried out in two sessions due to the length of some of the experiments. The total time for each study was 50 - 60 minutes. The administration time for each task in the pilot and the main study is set out in Table 5.1. All the experiments were administered on a one-to-one basis by a native Chinese speaker.

Table 5-1 The administration time for the tasks conducted in the pilot and the main study

Tasks Administered	Pilot Study	Main Study
Study 1 Visual discrimination of Chinese words	4 min.	4 min.
Study 2 Orthographic acceptability of pseudowords and nonwords	7 min.	10 min.
Study 3 Creative spelling in Chinese	15 min.	25 min.
Study 4 The phonological and semantic constraints in reading Chinese	10 min.	15 min.
Total	51 min.	54 min.

5.3 The Subjects

5.3.1 Sample for the pilot study

It was assumed that the period when Chinese children's knowledge of the Chinese written system grows more rapidly occur between four to nine years old. Thirty-six children, aged four to nine, were invited to undertake the experiments in the pilot study. The 4- and 5-year-olds were studying in kindergartens and the older children were in primary grades one to four. There were 21 boys and 15 girls altogether in this study. Table 5.2 presents the number of children in different age groups.

Table 5-2 Number of children in each age group in the pilot study

4-year-olds	2
5-year-olds	6
6-year-olds	6
7-year-olds	8
8-year-olds	8
9-year-olds	6

5.3.2 Sample for the main study

A larger sample was needed for further data analysis. The four-year-olds were excluded from the age range in the main study as their exposure to written Chinese was still limited. Altogether two hundred children, aged five to nine years participated in the main study. One hundred children took part in the experiment in Study 1 (Visual discrimination of Chinese words), and another hundred children were invited to undertake the series of experiments in Studies 2, 3 and 4. The age range, sex distribution and number of children in each age group in both samples were identical. There were twenty children, with equal numbers of boys and girls, in each age group. The children in each age group were approximately of the same age (plus or minus one month), with at least ten months difference between the age groups. These age groups

will be called hereafter 5-, 6-, 7-, 8-, and 9-year-olds. Table 5.3 presents the number of subjects, the sex distribution, and the children's ages.

Table 5-3 Age range, sex distribution and number of children in each age group in the main study

	Age Range (in months)	Sex Distribution		Number of Children
		Boys	Girls	
5-year-olds	59 -61	10	10	20
6-year-olds	71 - 73	10	10	20
7-year-olds	83 - 85	10	10	20
8-year-olds	95 - 97	10	10	20
9-year-olds	107-109	10	10	20

All the children were chosen from kindergartens and primary schools in middle class neighborhoods. In Hong Kong, it being a newly developed city, most of the people belong to the middle class. The parents have steady jobs, a reasonable income and secondary education or above.

5.3.3 Educational background of the subjects

Although formal education begins at the age of six in Hong Kong, over 98% of the children are sent to either nurseries or kindergartens at an earlier age. The informal teaching of reading and writing Chinese begins at the preschool level when children are as young as three. It is understandable that parents are anxious to let their children learn as early as possible; however, whether the teaching of written Chinese, especially the writing of Chinese, should be taught at such an early stage is still debatable.

Children at the age of five are expected to read about a hundred characters and to write ten to twenty characters. At the age of six, around two hundred characters are introduced in reading and about fifty to sixty for writing. Thus children beginning their

formal education are supposed to be equipped with a vocabulary of two hundred words, although a wide range of performance amongst them is expected. At primary level, it is stated in the “Syllabus for the Teaching of Chinese in Primary Schools” that children have to learn to recognize, and to a large extent to write about five hundred Chinese characters by Grade 1 (seven-year-olds), a thousand characters by Grade 2 (8-year-olds) and 1,500 characters by Grade 3 (nine-year-olds). By the end of the primary school years, children are expected to read and write about 2,500 characters.

Teaching of Chinese usually begins with the teaching of individual characters at preschool level. Characters with simple strokes are introduced first and are then followed by the more complicated characters. Two and three-character words are also introduced to preschool children. At primary level, children learn vocabulary from passages instead of isolated words. Textbooks recommended by the Education Department are usually adopted by the primary schools, and guidelines for teaching are available from the “Syllabus for the Teaching of Chinese in Primary Schools”. With regard to the teaching of vocabulary, the guidelines for teaching are fairly vague. The only requirement is that “children should be taught the shape, the meaning, and the pronunciation of the words”. It is unclear if the shape refers to the graphic features, the composition of the stroke-patterns, or merely how the various strokes occur in the word. Furthermore, there is no mention in any of the documents produced by the Education Department that the semantic radicals can be utilized to convey meaning and that the phonological component may give a clue to pronunciation, in analyzing individual characters. In other words, it is assumed that Chinese characters are learnt by rote. Both adults (parents and teachers) and children complain about the stress-inducing learning process for Chinese characters, and if the findings of the present study can offer some ways to improve this situation, it would be of great benefit to both adults and children.

Chapter 6

Study 1: Visual discrimination of Chinese words

6.1 *The Pilot Study*

6.1.1 What are the units of visual discrimination in Chinese?

It is commonly believed that visual complexity of a Chinese character is measured by the number of strokes (Cheng, 1992; Seidenberg, 1985; Hue and Erickson, 1988). A stroke is the smallest graphic unit in written Chinese, and a character consists of from one to over twenty strokes. All the strokes are packed in a confined two-dimensional square-shaped composition. Researchers who regard strokes as the unit of visual complexity assume that a word with more strokes is more difficult to identify than a word with fewer strokes.

This view, however, has been challenged recently by Fang & Wu (1989) and Chen (1993). They suggest that a stroke is the perceptual unit of simple words only, and for over 90% of the compound words, stroke-patterns are used as the basic perceptual units. Most of the compound words consist of two to three stroke-patterns and the more complicated words consist of four to six stroke-patterns. These are familiar graphic patterns and most of the time occupy fixed positions, although the size may differ in order to fit into the constricted square-shaped composition. The recursive combination of stroke-patterns in forming different words was demonstrated in Figure 2.2 in Chapter 2. Chen and Allport (1995) provide evidence that adult Chinese readers depend on stroke-patterns, instead of strokes, in their visual analysis of Chinese compound words. The speed of simultaneous “Same - Different” comparisons of Chinese words was shown to be affected by the number of stroke-patterns and by the proportion of mismatching units, but not by the number of individual strokes. They

concluded that stroke-patterns, rather than strokes, function as the constituent orthographic units of Chinese words for fluent adult readers.

In this study, we investigated whether stroke-patterns are used by children in visual discrimination of Chinese words. No study has ever investigated the basic perceptual unit being utilized by Chinese children learning to read. It is also noted that although strokes are explicitly and systematically taught in pre-schools when children begin to write, this is not the case for stroke-patterns. Some teachers may introduce some familiar stroke-patterns but they are never introduced systematically. In other words, using stroke-patterns as units of visual discrimination is a kind of implicit knowledge which is not learned through teaching.

In the pilot study, 20 pairs of Chinese pseudowords, which varied in the number of strokes and the number of stroke-patterns, were created. Children were asked to decide if the pairs of pseudowords were the same or different. If strokes are the units of visual processing of words, then children will probably make more mistakes in comparing pairs with more strokes than with fewer strokes. If stroke-patterns are the units of visual processing, then the degree of similarity of the stroke-patterns should affect the comparison.

6.1.2 Subjects

Thirty-six children, aged 4 - 9 participated in this study.

6.1.3 Design

Twenty pairs of pseudowords were used as stimuli for this experiment (Appendix 1), the pairs being devised to avoid the frequency effect of real words. Pseudowords were formed by substituting one stroke-pattern from another word. The pseudowords

looked like real words because all the stroke-patterns were in their conventional positions. The pairs of pseudowords differed in stroke-complexity and unit-complexity (Figure 6.1). Two groups of different stroke-complexity were devised - fewer strokes (5-12 strokes; average 8.9 strokes) and more strokes (13-20 strokes; average 13.4 strokes). Ten pairs had fewer strokes and the other ten pairs had more strokes. The twenty pairs of pseudowords also differed in their unit-complexity. There were five levels of unit-complexity: same, 2/3 same, 1/2 same, 1/3 same and totally different. In the 'same' pair, all the stroke-patterns in the pseudowords were identical. In the 2/3 same pair, one stroke-pattern was different while the other two stroke-patterns were the same; in the 1/2 same pair, one stroke-pattern was different; in the 1/3 same pair, only one stroke-pattern was identical while the others were different; and in the 'different' pair, no stroke-patterns were the same.

In addition to the pseudoword pairs, four sets of picture cards were prepared as a warming up exercise. This helped to ensure that the children could understand the concept of "same" and "different", and encouraged them to look at minor details on the card. Each pair of picture cards contained the same object (e.g. a car). The object might be slightly different (e.g. on one card the car has a bumper, and on another card the car has no bumper). Three pairs of picture cards were different from each other and the fourth pair was the same.

6.1.4 Materials

There were 20 white cards (24 cm x 7 cm) altogether. On each card, a pair of pseudowords was shown; the pseudowords were separated from one another by 8 cm. The size of a pseudoword was 4 cm x 4 cm.

Figure 6.1 Distribution of the 20 pairs of pseudowords by unit-complexity and stroke-complexity in Study 1

Unit-complexity	Stroke-complexity	Examples
Same (4 pairs)	Fewer Strokes (2 pairs)	𠂇 𠂇
	More Strokes (2 pairs)	黔 黔
2/3 Same (4 pairs)	Fewer Strokes (2 pairs)	𠂇 𠂇
	More strokes (2 pairs)	𠂇 𠂇
1/2 Same (4 pairs)	Fewer Strokes (2 pairs)	𠂇 𠂇
	More Strokes (2 pairs)	𠂇 𠂇
1/3 Same (4 pairs)	Fewer Strokes (2 pairs)	𠂇 𠂇
	More Strokes (2 pairs)	𠂇 𠂇
Different (4 pairs)	Fewer Strokes (2 pairs)	𠂇 𠂇
	More Strokes (2 pairs)	𠂇 𠂇

6.1.5 Procedure

The experiment was conducted individually. The experimenter invited the child to sit next to her, and the following instructions were given:

“Are the pair of pictures the same or different?”

“How do you know that they are the same/different?”

“Are the pair of words the same or different?”

The cards were shuffled for every trial. They were then shown to the child one at a time, and were then followed by another card. The child was required to say whether the words were the same or different. The experimenter noted their answers on a record sheet.

6.1.6 Results

6.1.6.1 *Stroke-complexity*

The mean scores for correct judgement of different stroke-complexity by age groups is presented in Table 6.1. There were ten pairs of pseudowords altogether with fewer strokes (5 - 12 strokes) and ten pairs with more strokes (13 - 20 strokes). The degree of stroke-complexity had very little effect on correct identification of the word pairs for all age groups. Even the 4-year-olds could get 80 percent correct for the more complex stroke type.

Table 6-1 Mean scores of correct judgement in different stroke-complexity by age group

Age	5 -12 strokes	13 -20 strokes
4-year-olds (n=2)	10	8.0
5-year-olds (n=6)	9.6	9.6
6-year-olds (n=6)	9.8	9.8
7-year-olds (n=8)	10	9.8
8-year-olds (n=8)	10	9.5
9-year-olds (n=6)	10	9.6

A mixed design ANOVA was used to compare the performance of children of different age groups in making correct judgements in different stroke-complexity. The factors considered were age as a between factor and stroke-complexity (more strokes and fewer strokes) as a within factor. No effect of age was found $F(5,30)=0.46$, but

there was a significant stroke-complexity effect, $F(1,30)=22.84$, $p<0.001$ and a significant age by stroke-complexity interaction effect, $F(5,30)=5.48$, $p<0.05$. The significant effect of stroke-complexity resulted from the performance of the four-year-olds who did relatively badly in differentiating the pairs with more strokes. When the four-year-olds were taken out from the analysis, then the stroke-complexity effect was no longer significant, $F(1,29)=2.52$, and the interaction effect disappeared, $F(4,29)=0.52$.

6.1.6.2 Unit-complexity

Table 6.2 presents the mean scores of different age groups in making correct judgements in various unit-complexity categories. There were four pairs of pseudowords in each category: same; 2/3 same; 1/2 same; 1/3 same; and different.

Table 6-2 Mean scores of correct judgement in different unit-complexity by age group

Age	same	2/3 same	1/2 same	1/3 same	different
4-year-olds (n=2)	3.5	3	3.5	4	4
5-year-olds (n=6)	3.3	4	4	4	4
6-year-olds (n=6)	3.8	3.8	4	4	4
7-year-olds (n=8)	4	3.8	4	4	4
8-year-olds (n=8)	4	3.7	4	4	4
9-year-olds (n=6)	4	3.6	4	4	4

A repeated measure ANOVA was applied to look at the effect of different unit-complexity by age groups. The factors considered were age as a between factor and the five levels of unit-complexity as a within factor. There was no age effect, $F(5,30)=0.96$ but a significant unit-complexity effect was found, $F(4,120)=4.38$, $p<0.001$. The interaction effect between age and unit-complexity was not significant, $F(20,120)=1.04$.

A significant unit-complexity effect was found because children made more mistakes in the '2/3 same' category. The pair of pseudowords in '2/3 same' category were very similar. There were altogether four '2/3 same' pairs, two pairs with fewer strokes (6-12 strokes) and two pairs with more strokes (13 - 20 strokes), and all the mistakes occurred on the pairs with more strokes. Since the two pairs of '2/3 same with more strokes' were the only ones that made a difference in children's performance, the distribution of errors within these two pairs will be discussed in more detail. One pair of stimuli [樹₍₁₃₎ 樹₍₁₅₎] has three stroke-patterns with fewer strokes and the other pair [足₍₁₇₎ 足₍₂₀₎] have two stroke-patterns with more strokes. If strokes, instead of stroke-patterns are the units of visual analysis, then we would expect the errors to occur on the pair with fewer stroke-patterns but with more strokes. On the other hand, if stroke-patterns are the units of visual analysis, then the pairs with more stroke-patterns are likely to lead to more mistakes, even though they have fewer strokes. Results from this experiment showed that all the children, except the four-year-olds, made mistakes in the pair with three stroke-patterns with fewer strokes, instead of the pair with two stroke-patterns with more strokes.

Apart from the '2/3 same' category, some younger children made some mistakes in judging the 'same' pairs. They tended to say that the two words in the 'same' pairs were different because the length of the strokes was not identical. They were distracted by the minor flaws of the test instruments because the pair of words were written instead of photocopied. It also demonstrated that the 4-year-olds have not yet developed an understanding that the length of the strokes does not affect the orthographic structure of the words.

6.1.7 Conclusions

Most of the children in this study could differentiate Chinese words of various stroke-complexity and unit-complexity without much difficulty. It is really surprising that the

* Number of strokes in the word

4-year-olds can do so well as they have not yet been formally taught to read. The frequent exposure to environmental print is probably one of the main reasons for the early development of visual skills for word discrimination. The performance of the 6-, 7-, 8- and 9-year-olds almost reached ceiling, thus there was no age difference in their ability to differentiate the word pairs.

This experiment failed, to a large extent, to examine the strategies children use in visual discrimination because the task was too simple for them. The pseudowords were too big (4 cm x 4 cm) to allow mistakes to occur. Most of the children had no difficulty differentiating the 'same' pairs irrespective of the number of strokes and stroke-patterns. They had no difficulty differentiating pseudowords with half or more than half different, as the difference was too obvious and they could compare it on a whole word basis. The only two pairs of pseudowords that demanded acute visual discrimination were the '2/3 same pairs with more strokes' and it was found that three-stroke-pattern words were more difficult to judge than the two-stroke-pattern words, although the two-stroke-pattern words had more strokes. Further evidence is still needed as the performance of the older children had reached ceiling and evidence of using stroke-patterns rather than strokes was gathered from a limited number of children. In the main study, by introducing word pairs which demanded acute discrimination, the confounding factors were modified. Pairs of stimuli would vary only in limited number of strokes and stroke-patterns. In addition, equal numbers of same and different pairs would be used, sixteen out of twenty pairs of stimuli in the pilot study were different and children had a tendency to say 'no' to the stimuli. Furthermore, the stimuli for the 'same' pairs would be photocopied words to avoid children being distracted by inconsistency in writing the same word.

6.2 The Main Study

6.2.1 Difficulty of visual discrimination: number of strokes versus number of stroke-patterns

The experiment was re-designed to investigate again if children use strokes or stroke-patterns as the visual unit in discriminating words. A new set of stimuli was devised and children were required to judge whether a pair of pseudowords was the same or different.

6.2.2 Subjects

One hundred children, age 5 - 9, participated in this study (for a fuller description, see Section 5.3.2 in Chapter 5).

6.2.3 Design

There were altogether ten identical pairs and ten different pairs of pseudowords (Appendix 2). In the identical pairs, four of them consisted of 8 - 9 strokes and six of them consisted of 14 - 19 strokes. At the same time half of them consisted of two stroke-patterns and half of them three stroke-patterns, as shown in Figure 6.2. It was predicted that if strokes are the unit of visual analysis, more mistakes would be found in discriminating the pairs with more strokes as it would be more difficult to compare words with more strokes than with fewer strokes. On the other hand, if the stroke-pattern is the unit of visual analysis, then pairs of words with more stroke-patterns would be more difficult to discriminate than words with fewer stroke-patterns, irrespective of the number of strokes in the words.

Figure 6.2 Distribution of the ‘same’ pairs of pseudowords by stroke- and unit-complexity

	2-stroke-patterns		3-stroke-patterns	
	Number of pairs	Examples	Number of pairs	Examples
fewer strokes (8 - 9 strokes)	2	佻 佻	2	廔 廔
more strokes (14-19 strokes)	3	稿 稿	3	蛸 蛸

In the ‘different’ pairs, all the words consisted of 12 - 16 strokes, and each pair of words differed from one another by one to five strokes. At the same time, half of them consisted of two-stroke-patterns and half of them had three-stroke-patterns. The number of strokes that differed between the pair of words corresponded in the two- and three stroke-patterned pairs (Figure 6.3). If stroke-patterns are not utilized in reading Chinese, we expect a similar accuracy rate in discriminating pairs of words with the same number of stroke variation and total number of strokes, irrespective of the number of stroke-patterns. If stroke-patterns play a role in reading Chinese, then 3-stroke-patterned pairs would result in more mistakes than 2-stroke-pattern pairs, even though the pairs of words have the same number of stroke variation and total number of strokes.

Figure 6.3 Distribution of the ‘different’ pairs of pseudowords by unit-complexity and number of stroke difference

	2 stroke-patterns	3-stroke-patterns
1 stroke difference	靚 (16)* 靚 (15)	瘰 (16) 瘰 (15)
2 strokes difference	霄 (13) 霄 (13)	踞 (14) 踞 (14)
3 strokes difference	輶 (14) 輶 (14)	覷 (14) 覷 (14)
4 strokes difference	魚 (14) 魚 (14)	荳 (13) 荳 (15)
5 strokes difference	詆 (12) 詆 (12)	矧 (12) 矧 (12)

* number of strokes in the words

The stroke difference is shown here by a different colour but they were all black when presented to subjects

6.2.4 Materials

The twenty pairs of pseudowords were written in a 2 cm x 2 cm square. Each pair of pseudowords were placed on a purple card 15 cm x 4 cm and they were separated from one another by 5 cm.

6.2.5 Procedure

The experiment was conducted in the form of a competition between two children. They were asked to discriminate as fast as they could twenty pairs of pseudowords. They had to put the 'same' pairs on a big yellow card with the word 'SAME' written on it, and the 'different' pairs on another yellow card with the word 'DIFFERENT' written on it. The first person to finish reading all the words was required to run and get a red ball and the second person to finish picked up a yellow ball. In this way, children were encouraged to judge the words as quickly as possible. Two examples with real word pairs were given before the experimental trials, one pair of words was the same and the other pair was different.

The following instructions were given:

For the two examples,

"There are two words written on this paper, are they the same or are they different?"

If the children could differentiate the examples, they proceeded to the experimental trials. If they could not differentiate the examples, then they would not proceed to the experiment. All the children could differentiate the examples.

For the experimental trials,

"There are twenty pairs of very difficult words and you are going to judge if they are the same or different. If they are the same, put the card on this paper with the word SAME on it, if they are different, put the card on the piece of paper with the word DIFFERENT on it. You have to do it as quickly as possible. The first person who finishes reading all the words will then run to the back of the classroom and get me a red ball, and the second one can only get me a yellow ball."

"Are you ready? One, two, three, start!"

6.2.6 Results

The distribution of scores in same and different pairs is shown in Appendix 3.

6.2.6.1 Stroke-complexity in 'same' pairs

In the 'same' pairs, the number of strokes and the number of stroke-patterns were varied. If the children used strokes as the unit of visual analysis, more mistakes would be found in the pairs with more strokes. On the other hand, if stroke-patterns are the unit of visual analysis, then more mistakes were likely to occur in the pairs with more stroke-patterns. The mean and standard deviation of correct judgment in the fewer strokes pairs and more strokes pairs are presented in Table 6.3. The degree of stroke-complexity had very little effect on correct identification of the word pairs for all age groups. Most of the children identified the 'same' pairs correctly and a ceiling effect was obtained.

Table 6-3 Mean and standard deviation of the number of correct judgements at different levels of stroke-complexity by age group

Age level	Stroke-complexity	
	fewer strokes (8 - 9 strokes)	more strokes (14 - 19 strokes)
5-year-olds	0.95 (0.13)	0.97 (0.12)
6-year-olds	0.98 (0.11)	0.92 (0.14)
7-year-olds	0.96 (0.09)	0.96 (0.15)
8-year-olds	0.99 (0.56)	0.97 (0.07)
9-year-olds	1.0	0.96 (0.12)

Table 6.4 presents the mean and standard deviation of the number of correct judgments for different levels of unit-complexity by age group. The degree of unit-complexity also had very little effect on correct identification of the word pairs for all age groups.

Table 6-4 Mean and standard deviation of correct judgement in different unit-complexity by age group

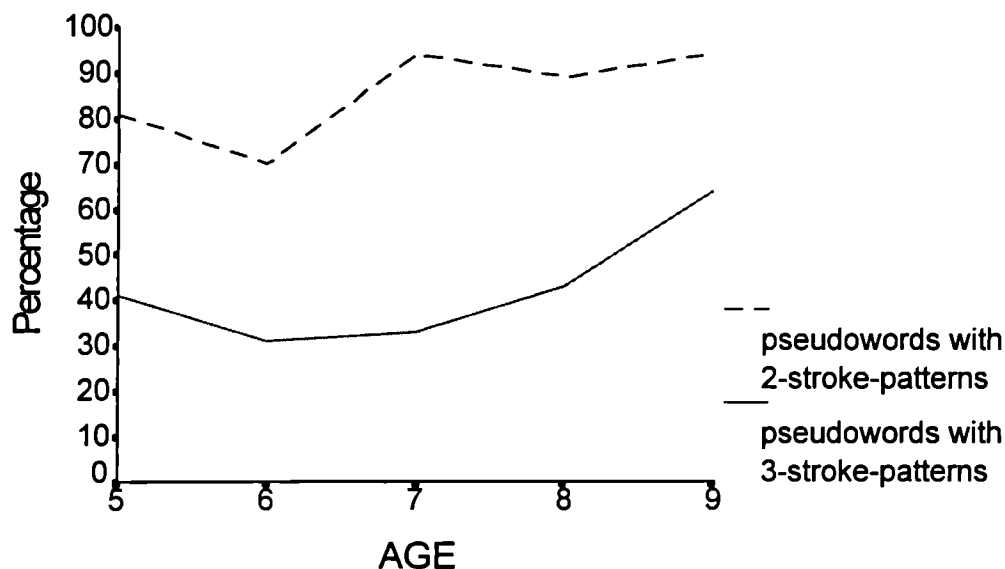
Age level	Unit complexity	
	2-stroke-patterns	3-stroke-patterns
5-year-olds	0.95 (0.11)	0.97 (0.13)
6-year-olds	0.95 (0.14)	0.93 (0.12)
7-year-olds	0.96 (0.10)	0.96 (0.10)
8-year-olds	0.98 (0.09)	0.97 (0.73)
9-year-olds	0.98 (0.62)	0.98 (0.62)

A mixed design ANOVA was conducted to compare the performance of children in making correct judgments in more complex stroke-pairs (14 - 19 strokes) and more complex unit-pairs (3-stroke-patterns). The factors considered were age as a between subjects factor and the two kinds of complexity (stroke vs. unit) as a within subjects factor. No significant age effect was found, $F(4,95)=0.73$, there was no difference in the correct judgments of stroke- and unit-complexity, $F(1,95)=1.66$, and there was no interaction between age and the correct judgments in the two kinds of complexity, $F(4,95)=0.23$.

6.2.6.2 Unit-complexity in 'different' pairs

Figure 6.4 presents the percentage of correct responses in the 2-stroke-pattern pairs and the 3-stroke-pattern pairs. Children of all age groups made more mistakes in identifying 'different' pairs with 3-stroke-patterns than with 2-stroke-patterns.

Figure 6.4 Percentage of correct responses for 'different' pairs with 2- and 3-stroke-patterns by age groups



A mixed model ANOVA was used to compare the performance of children in making correct judgments at different levels of unit-complexity. The factors considered were age as a between subjects factor and unit-complexity (2 stroke-patterns and 3 stroke-patterns) as a within subjects factor. A significant age effect, $F(4,95)=6.09$, $p<0.001$ and a significant unit-complexity effect, $F(1,95)=214.02$, $p<0.001$, and a significant age by unit-complexity interaction effect, $F(4,95)=3.02$, $p<0.05$ were found. The interaction effect was caused by the gradual improvement in children recognizing three-patterned words and their relatively inconsistent performance in discriminating two-patterned words. Post hoc examination by Scheffe test revealed that the six-year-olds were less capable of discriminating the word pairs with 2-stroke-patterns than the 7-, 8- and 9-year-olds. Post hoc examination by Scheffe test also showed that 6- and 7-year-olds were significantly worse at differentiating the word pairs with 3-stroke-patterns than the nine-year-olds.

6.2.7 Conclusion

The purpose of this study was to investigate if stroke-patterns are used by children as a perceptual unit. A ceiling effect was found in discriminating the 'same' pairs. Children in all age groups could differentiate over 90% of the words correctly, and no age effect was found. The result confirms the findings in the pilot study that very young children have already developed sophisticated strategies in recognizing the orthographic features of Chinese words. Older children spent about three seconds and five-year-olds spent about six seconds to differentiate a pair of words. No child has demonstrated an inability to make systematic comparison of the orthographic features of the Chinese words.

The reason which accounts for the sophisticated word recognition ability in Chinese children is unlikely to be related to global visual skills. Many studies (Lee, Stigler & Stevenson, 1986; Huang & Hanley, 1992; Chen, Lau and Yung, 1993) have already demonstrated that Chinese children are not more advanced in global visual skills than alphabetic readers and that the relationship between general visual skills and reading Chinese is not strong. In an attempt to discover the precursors of learning to read Chinese, Ho (1994) conducted extensive visual skills testing with a group of Chinese preschool children and she found that the ability to detect 'constancy of shape' is the only visual skill that is a significant predictor of children's Chinese word reading scores. Ho suggests that it is the particular characteristics of Chinese orthography, the recurring nature of stroke-patterns, that facilitates Chinese readers to recognize the words quickly and accurately: "the same stroke-pattern, though recurring, can appear in different sizes at different location of different characters..... the ability to detect constancy of shape as an important predictor of beginning reading success in Chinese seems to reflect the importance of detection of stroke-patterns in Chinese character recognition" (p.189)

Due to the ceiling effect, no effect of either stroke-complexity or unit-complexity effect was found in the 'same' pairs. However, in the 'different' pairs, the level of

difficulty was elevated by introducing slight differences between the pair of words. The task demanded acute discrimination because the word pairs were different in one to five strokes only. The total number of strokes was controlled and only complicated pseudowords were used. With the increase in difficulty, children made more errors and they could only get 64% correct in discriminating 'different' pairs. A significant age effect was found with older children being able to differentiate word pairs more accurately than young children.

Furthermore, a marked difference was found in discriminating word pairs with different unit-complexity. While children could discriminate 86% of the word pairs with 2-stroke-patterns, they could only compare correctly 42% of the word pairs with 3-stroke-patterns. Since the number of stroke-patterns is the only factor that varied, while the total number of strokes in the words and the number of different strokes were controlled, it is thus concluded that unit-complexity plays a significant role in word recognition.

To conclude, the results from this experiment demonstrate that very young children have already acquired the visual skills needed in word recognition. It is very likely that young children, similar to the adults in Chen, Allport & Marshall's (in press) study, have already learned to use stroke-patterns as the basic unit of visual analysis because it is found that unit-complexity causes significant difference in discriminating words with minor differences. In the following studies, we examine children's understanding of the formal and functional constraints of these stroke-patterns. If evidence can be gathered demonstrating that children do not merely use the stroke-patterns for visual analysis, but extends its usage to representational functions, then these stroke-patterns are very likely to be the basic units of the Chinese written language system.

Chapter 7

Study 2: Orthographic acceptability of pseudowords and nonwords

7.1 *The Pilot Study*

7.1.1 The formal constraints of written Chinese - the positioning of stroke-patterns

In order to read an alphabetically-written word one must recognize the letters. Letter-mediated accounts are characteristic of most current theoretical frameworks for word recognition in English orthography (e.g. Colheart, 1981; Henderson, 1982; Rayner & Pollatsek, 1989). It has been shown in several independent investigations that, before children learn to read and write an alphabetic script, they have already developed some understanding of the formal constraints of their written language. This understanding works as a basis for learning to read and is predictive of children's reading progress (Ferreiro and Teberosky, 1982; Nunes Carraher and Rego, 1982). It continues to grow as the children learn more about the written language. Ferreiro and Teberosky (1982), for example, showed that preschool children in Argentina, who did not know how to read, used some definite criteria in order to decide whether a graphic pattern could or could not be read. They used a minimum quantity criterion, according to which at least three letters are necessary for reading, with single and pairs of letters not being enough for reading. Henderson & Chard (1980) conducted a lexical-decision experiment, in which nonwords are classified according to presence or absence of vowel and single letter positional frequency. They found that second grade children can use their analysis of positional redundancy to reject nonwords. Thus even six- to seven-year-olds are sensitive to quite subtle features of orthographic structure.

On the other hand, word recognition in Chinese is often regarded as holistic, with no mediation by its orthographic constituents. More recently, however, this supposition has been critically re-examined. Chen (1993) proposed the stroke-pattern to be the equivalent of the “letter” as the functional orthographic unit in Chinese. Stroke-patterns play a crucial role in word recognition because they are familiar visual patterns which appear as a constant spatial design and occupy a specific position. Putting a stroke-pattern in an illegal position would violate the orthographic structure and result in a nonword which would be rejected by adult readers (Chen, 1993). This view has recently been acknowledged by many psychologists who are interested in the cognitive processing of Chinese (Peng, Li & Huang, 1995, Cheng & Huang, 1995).

In contrast to alphabetic words, words in Chinese are not conventionally spelled by a set of components like the letters of English. Explicit orthographic knowledge of letters is universal among competent readers of English. The constituents of composite Chinese characters are not conventionally taught independent of word context. Nor are such orthographic units systematically classified by scholars as word constituents, with the exception of the set of components listed in the dictionary, which serves as its index key. Despite this, it is predicted that an understanding of the underlying orthographic rules will gradually develop among Chinese readers as a tool to help them to learn to read.

This study examined children’s ability to make orthographic decisions on the basis of the legal positioning of stroke-patterns. Two types of stimuli were used: pseudowords, where stroke-patterns are in their legal positions, and nonwords, where stroke-patterns are in an illegal position. If children reject all the stimuli, then a strong lexical effect is demonstrated, that is, they will only accept real words. If children accept all the words, it means that they have no underlying orthographic rules in deciding what a word is, and any ‘Chinese-like’ graphic units can be regarded as a Chinese word. If children accept pseudowords and reject nonwords, it demonstrates that children have acquired the knowledge of legal positioning of stroke-patterns and rely on this rule to judge whether

any unknown words can be accepted as words. This pattern of performance would support the hypothesis that written Chinese is not learnt by rote, instead, an understanding of the orthographic structure is also acquired in the process of learning to read. Similar to English readers, Chinese learners also acquire rule-based orthographic structures in order to help them to read and write Chinese.

7.1.2 Subjects

Thirty-six children, aged 4 - 9 participated in this study.

7.1.3 Design

Ten pseudowords and ten nonwords were devised for this experiment (Appendix 4). As none of the stimuli are words, there is no word frequency effect. The only difference between pseudowords and nonwords is the positioning of stroke-patterns. Pseudowords were constructed by the substitution of one unit from real words to form a non-existing combination of orthographic units. Radicals of pseudowords are always in their legitimate positions. Nonwords were constructed by placing the radicals of words in illegal positions.

Figure 7.1 Examples of pseudowords and nonwords for the pilot study

Word Types	Stroke-patterns	Legal Positions	Examples
Pseudowords	丶 朱	left - right	朱
	冫 主	top - bottom	主
Nonwords	丶 可	left - right	可
	冫 牛	top - bottom	牛

7.1.4 Materials

There were three sheets of A4 size paper, with six to seven words written on each page. Each word was approximately 2 cm x 2 cm. The pseudowords and nonwords were randomly mixed and written on the three sheets of paper.

7.1.5 Procedure

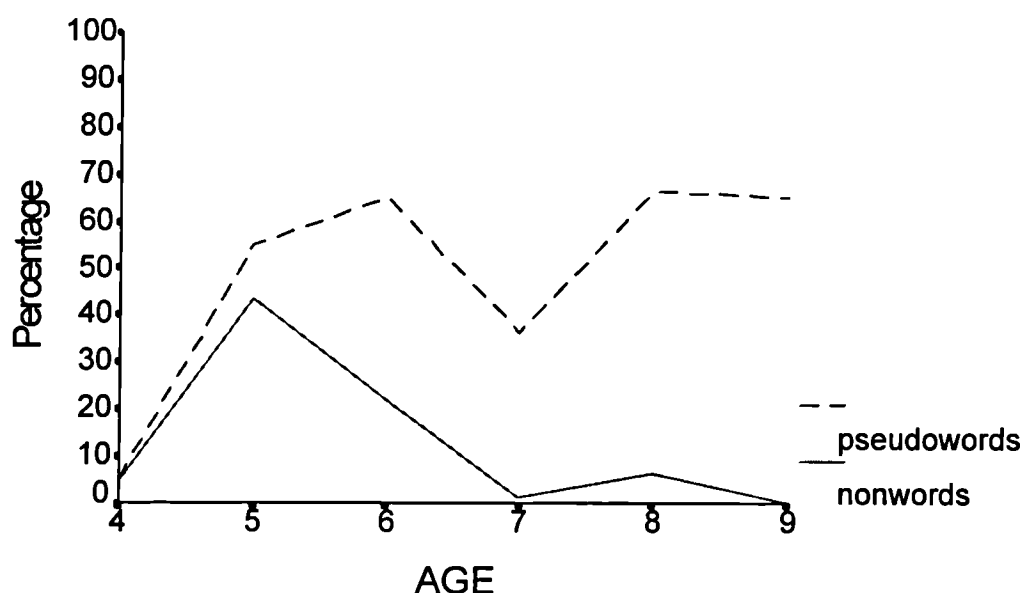
A toy monkey was introduced in this task. The children were told that the monkey had written some words on the sheets of paper. Sometimes he got it right but sometimes he got it wrong. The child was invited to be a 'teacher' and to mark the paper by putting a tick if the word was correct and a cross when it was wrong. The following instructions were given to each child:

“Pretend you are a teacher, here is a red pen. Please mark the papers. There are altogether three sheets of paper. Put a tick if the word is correct and a cross if it is wrong.”

7.1.6 Results

Figure 7.2 presents the percentage of children accepting the stimuli as words by age groups. The 4-year-olds had a tendency to reject both pseudowords and nonwords. Five-year-olds accepted more pseudowords and nonwords than the 4-year-olds, and there was a tendency to accept more pseudowords than nonwords already at this age level. The gap widened for older age groups, with the nine-year-olds being able to reject all the nonwords yet accept most of the pseudowords.

Figure 7.2 Percentage of orthographic acceptability of pseudowords and nonwords by age group



A mixed-model ANOVA was used to examine the effects of age and correctness of position of radicals on the judgments of orthographic acceptability as the dependent variable. There was a significant age effect, $F(5,28)=5.29$, $p<0.05$, a highly significant effect of orthographic acceptability, $F(1,28)=41.99$, $p<0.001$ and a significant age by positional regularity effect $F(5,28)=3.39$, $p<0.05$. Older children accepted pseudowords and rejected nonwords whereas young children could not make the differentiation so clearly.

7.1.7 Discussion

The experiment investigated if children have acquired the knowledge of the formal constraint of written Chinese, by referring to the positioning of the stroke-patterns. The impressive results indicated that children accept pseudowords and reject nonwords according to the legal positioning of the stroke-patterns. They reject the nonwords because the stroke-patterns are in illegal positions, while pseudowords can be accepted as the stroke-patterns are in their conventional positions. The results also demonstrate

that although the children do not know any of the pseudowords and nonwords, they follow consistent orthographic rules in accepting or rejecting the stimuli, on the basis of implicit knowledge. Only the four-year-olds do not demonstrate a preference for pseudowords over nonwords. There is no evidence that they have strategies to decide whether a stimulus is a word or not, and they tend to reject most of the stimuli.

It is rather surprising to see that the 5-year-olds already show a tendency to reject nonwords as they have only recently been introduced to informal teaching of written Chinese. The influence of environmental print may be one possible explanation. Goodman (1984) has pointed out that in a print-rich environment of most present cultures, young children are continuously interacting with, organising, and analysing the meaning of the visible language. The older children have begun to learn how to read and write and they are able to make sensible judgements - accepting pseudowords and rejecting nonwords.

A shortcoming of the study is the possibility of facilitating negative responses to nonwords because all the nonwords were made from words by shifting the positions of stroke-patterns. Children may have paid attention to the positioning of the stroke-patterns only because they discovered that the words were written incorrectly. As one child explicitly pointed out “this word is wrong, it is *river* < 𡿨𡿨 >, and this radical should be put on the right hand side, not the left hand side < 𡿨𡿨 >.”

7.2 The Main Study

7.2.1 The formal constraints of written Chinese - the graphic characteristics of the stroke-patterns

This study is a further investigation of young children’s acquisition of the knowledge of legal positioning and legal graphic design for stroke-patterns. The effect of positioning of stroke-patterns on orthographic acceptability had been tested in the pilot

study and this experiment was a replication study; the confounding factor in the pilot study was modified, and nonwords were constructed by placing radicals of pseudowords, instead of words, in illegal positions.

In addition, in order to increase the discrimination across categories of stimuli among the younger children, another kind of nonword with illegal graphic components, for example, elements which are not strokes in Chinese - circles, circular curves, etc., was added to this study.

Ferreiro and Teberosky (1982) found that preschool children consider both quantity and quality of graphic characteristics in alphabetic scripts. In their study, children seemed to recognize formal constraints of written language before they could read. For example, they rejected 'AAA' as something that can be read because 'this was just the same letters again'. They also rejected strings that contained elements which were not letters, and isolated letters.

A similar finding was obtained in Chan's (1990) study, when Chinese preschool children were asked to write some words that they had not yet learned. Older children were able to create approximate writing by using graphics similar to Chinese writing, such as vertical and horizontal strokes, while younger ones could only make scribbles. However, it is still not known in what ways Chinese children come to differentiate formal aspects of Chinese. Using illegal graphic patterns not designed specifically for Chinese writing would violate the orthographic structure. The strokes used as elements in stroke-patterns in written Chinese are limited. Circles, for example, are never used. Do children realize that some graphic forms are not strokes? Do they use this knowledge to decide on the orthographic acceptability of graphic patterns? No research has yet looked into the ability of children to recognize these kinds of graphic properties of Chinese words. In this study, nonwords with illegal graphic features

were included, as well as pseudowords and nonwords with illegal positions of stroke-patterns.

7.2.2 Subjects

One hundred children, aged 5 - 9 participated in this study. (for a fuller description, see Section 5.3.2 in Chapter 5).

7.2.3 Design

In this task, children were invited to be a 'teacher' and mark the words written by a child living on another planet who was trying to learn how to write Chinese. There were altogether thirty words, ten pseudowords and twenty nonwords (Appendix 5). Pseudowords were constructed by substitution of one unit from real characters to form a non-existing combination of graphic units. Lexical radicals of pseudowords were always in their legitimate positions. There were two types of nonwords: (1) they were constructed by placing the lexical radicals of pseudowords in illegal position.; or (2) they were made by substituting illegal graphic features in pseudowords (Figure 7.3).

Figure 7.3 Examples of pseudowords and nonwords for the main study

Type of words	Examples
Pseudowords with correct positioning of stroke-patterns	石已 侑
Nonwords with illegal positioning of stroke-patterns	平石 勿
Nonwords with illegal graphic features (a) circles, (b) multiple curves (c) illegal overall shapes (d) wrong stroke direction (e) wrong proportion	(a) 咄 (b) 泣 (c) 田 (d) 弓 (e) 𠂇

— semantic radicals — illegal graphic designs

7.2.4 Materials

There were altogether three sheets of A4 size paper, with ten words written on each page. Each word was approximately 2 cm x 2 cm. The pseudowords and nonwords were randomly mixed and written on the three sheets of paper.

7.2.5 Procedure

The following instruction was given to each child:

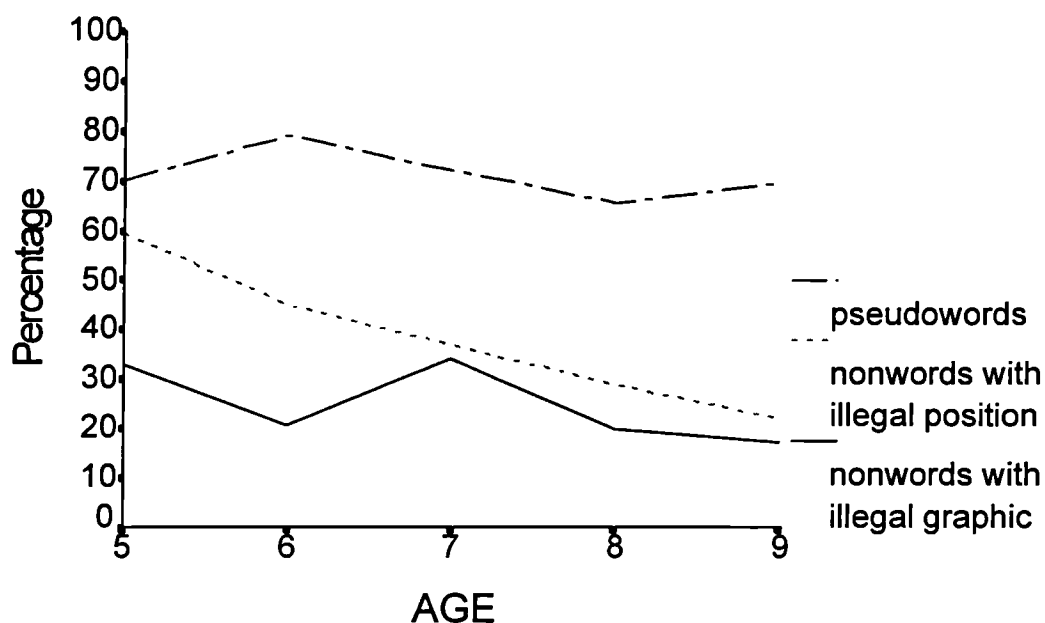
“Pretend you are a teacher, here is a red pen. Please mark the papers for the spaceman. Put a tick if the word is correct and a cross if it is wrong.”

7.2.6 Results

The distribution of scores of pseudowords, nonwords with illegal positioning of stroke-patterns, and nonwords with illegal graphic features is shown in Appendix 6. Children of all age groups accepted pseudowords (71%) more than nonwords with illegal positioning of stroke-patterns (39%) and in turn, accepted nonwords with illegal

positioning of stroke-patterns more than nonwords with illegal graphic features (25%). Figure 7.4 presents the percentage of children accepting the stimuli as words by age groups.

Figure 7.4 Percentage of orthographic acceptability of pseudowords and nonwords by age groups



One score was given to positive responses to each item. Thus a child who regarded three pseudowords, seven nonwords with illegal positioning of stroke-patterns as 'words' and rejected all nonwords with illegal graphic features scored three points for pseudowords, seven points for nonwords with illegal positioning of stroke-patterns and no points for nonwords with illegal positioning of stroke-patterns. A mixed-model ANOVA was conducted with the response to orthographic acceptability effect in the three types of stimuli as a within factor, and age as a between factor. There was a significant age effect, $F(4,95)=4.05$, $p<0.05$, a significant orthographic acceptability effect, $F(2,190)=135.96$, $p<0.001$ and a significant age by orthographic acceptability effect, $F(9,190)=3.14$, $p<0.05$. Older children responded differently as compared to younger children, especially in accepting 'nonwords with illegal positions' as 'words', and a downward

trend could be traced. In spite of this, even five-year-olds accepted more pseudowords than the two types of nonwords, and the preference remained the same across the age groups. The interaction effect was caused by the widening of the gap in orthographic acceptability of pseudowords and nonwords at different age levels.

The mean for positive responses to pseudowords and ‘nonwords with illegal positions’ is presented in Table 7.1 by age groups. A series of T-tests were conducted and the mean difference was significant in the 6-, 7-, 8- and 9-year-olds. A one-way ANOVA was conducted to look at the age effect with the mean difference between the responses to pseudowords and ‘nonwords with illegal positioning of stroke-patterns’ as a dependent variable. A significant age effect was found, $F(4,99)=5.85$, $p<0.001$. Post hoc comparison by Tukey-HSD tests revealed that the 5-year-olds were less capable of differentiating the nonwords from pseudowords than the older children.

Table7-1 Mean of positive responses to pseudowords and ‘nonwords with illegal positions’ by age groups

	Mean (SD)		τ	df	P
	pseudowords	nonwords with illegal positioning			
5-year-olds	7.0 (2.7)	6.0 (2.0)	2.06	19	0.05
6-year-olds	7.9 (2.0)	4.5 (2.5)	8.64	19	0.00*
7-year-olds	7.2 (2.7)	3.7 (2.3)	4.67	19	0.00*
8-year-olds	6.6 (2.8)	2.9 (2.5)	6.32	19	0.00*
9-year-olds	7.0 (2.5)	2.2 (2.1)	9.37	19	0.00*

* $p<0.001$

Table 7.2 presents the mean of positive responses to pseudowords and ‘nonwords with illegal graphics’ by age groups. The mean difference was significant in all age groups. A one-way ANOVA was also conducted to look at the age effect with the mean difference between the responses to pseudowords and ‘nonwords with illegal features’ as a

dependent variable. No age effect was found, $F(4,99)=1.49$. The gap between the pseudowords and nonwords with illegal graphic features did not vary significantly with age.

Table 7-2 Mean of positive responses to pseudowords and ‘nonwords with illegal graphics’ by age groups

	Mean (SD)		τ	df	P
	pseudowords	nonwords with illegal graphics			
5-year-olds	7.0 (2.7)	3.3 (2.5)	4.23	19	0.00*
6-year-olds	7.9 (2.0)	2.1 (1.7)	10.56	19	0.00*
7-year-olds	7.2 (2.7)	3.4 (2.7)	3.89	19	0.00*
8-year-olds	6.6 (2.8)	2.0 (2.3)	7.27	19	0.00*
9-year-olds	7.0 (2.5)	1.7 (1.7)	7.76	19	0.00*

* $p<0.001$

In recognizing nonwords with illegal graphic features, the percentage of accepting the following graphic features as words were: 14% for ‘nonwords with circles’, 11% for ‘nonwords with multiple curves’, 13% for ‘nonwords with illegal shapes’ and 22% for ‘nonwords with inproportional sizes of stroke-patterns’.

7.2.7 Discussion

In the pilot study, nonwords were made by shifting the stroke-patterns of real words. Children might have recognized that a nonword in the pilot study was wrong because they recognized the real word. In the main study, this confounding factor was avoided. Nonwords in the main study were formed by placing the stroke-patterns of pseudowords in illegal position. It was found that children in the main study accepted more pseudowords as ‘words’ (71%) than children in the pilot study (59%). At the same time, children in the main study accepted more nonwords with illegal positioning of stroke-patterns (39%) than children in the pilot study (6.6%). Although children

mistook more nonwords with illegal positioning of stroke-patterns as words in the main study, the age trend with older children accepting pseudowords and rejecting nonwords was similar to the one in the pilot study. The significant age difference found in systematically accepting pseudowords and rejecting nonwords in the main study provides further evidence in support of this claim.

The results converge with those of a recent study carried out by Cheng and Huang (1995) in mainland China. They recently reported a study examining the knowledge of the positioning of stroke-patterns of school children from Grade 2 to Grade 6. Children in mainland China are exposed to the simplified version of Chinese characters instead of the complicated version; in spite of this difference in type of character used, children in both studies accepted pseudowords and rejected nonwords as 'words', on the basis of legal positioning of stroke-patterns.

A new type of stimuli - nonwords with illegal graphic features - was added to the main study. Most of the children were able to reject them as words. Even the youngest age group, the five-year-olds, accepted only 33% of illegal graphic nonwords while they acknowledged 60% of 'nonwords with illegal positioning of stroke-patterns' and 70% of pseudowords as words. The five-year-olds had been introduced to reading a short time prior to the experiment, and they were not expected to know more than one hundred characters. However, they were able to reject stimuli which contained illegal graphic features. The illegal graphic features included in this experiment are not extraordinary. In fact, some of them can be found in artistic work, for example, changing the dots into circles (type A), or putting the multiple curves into waving lines (type B) for 'eye-catching' purposes. Words with wrong proportions (type E) can commonly be found in young children's writing when their fingers cannot yet control the length of the strokes, and their ability to orientate the stroke-patterns spatially is not yet fully developed. On the other hand, many children regarded 'nonwords with wrong stroke directions' (type D) as 'words', this accounted for 40% of the total errors

in the category 'nonwords with illegal graphic features'. It may be related to the mirror image writing that occurs quite often in young children.

Despite the difference between Chinese and alphabetic scripts, the results in this study are comparable to the findings in Lavine's (1977) and Ferreiro and Teberosky's (1982) studies where they investigated children's awareness of the formal constraints of their written language - that young children have precise ideas about the characteristics of written text even before they have learned how to read. The results are particularly significant because the formal graphic characteristics of written Chinese are not taught explicitly at any stage in school, thus the children must have an active role in the construction of their conception of formal graphic features of Chinese words.

To conclude, from the age of five years, children are sensitive to the formal constraints in the orthographic structure of Chinese. They utilize both their knowledge of graphic elements and legal positioning of the stroke-patterns to accept or reject stimuli as words. The result supports the hypothesis that children do not learn to read by rote; they acquire an understanding of the rule-based orthographic structure of written Chinese in the process of learning to read.

Chapter 8

Study 3: Creative spelling in Chinese

8.1 The Pilot Study

8.1.1 Representational functions of semantic radicals and phonological components

In Study 2, it was found that children can rely on the knowledge of positional and graphic constraints of written Chinese to decide the orthographic acceptability of pseudowords and nonwords. It is however, still unknown if they can impose these kinds of constraints in their spelling. In this study, children were asked to create some spellings with a given set of stroke-patterns. If children follow the orthographic rules, then they would create pseudowords in their spelling.

Furthermore, the stroke-patterns also have particular representational functions: they can either be semantic radicals or phonological components (Zhu, 1987). Studies on Chinese adult readers have provided evidence that these units are recognised and used in reading (Zhang & Simon, 1985; Hue & Erickson, 1988). Adult readers rely on the semantic radicals to identify the meaning of words and pseudowords (Flores d'Arcais, 1992; Chen, 1993) and on the phonological components for pronunciation (Seidenberg, 1985; Zhu, 1987; Hue, 1992).

The functions of stroke-patterns are indicated by their positions within a character. If a stroke-pattern constitutes a semantic radical, it has a fixed position in the pattern in any character where it is used with this function. Although the position may vary across semantic radicals, it is constant for each particular radical. Semantic radicals offer clues to

meaning; when serving this function, they are never pronounced. Some semantic radicals may be simple words on their own and have their own pronunciation as simple words; others are unpronounceable and do not appear on their own.

The remainder of the character is the phonological component, which also has a fixed position, and gives a clue to pronunciation. However, the pronunciation of a phonological component is not a simple issue: a phonological component may have the same pronunciation when serving this function as it has when it exists as a simple word on its own, or it may be pronounced differently when it is used as a phonological component.

This study was designed to investigate the following aspects of children's understanding of written Chinese: 1) the positioning of semantic radicals and phonological components; 2) the use of semantic radicals to express meaning; and 3) the reliance on phonological components to generate pronunciation. Three predictions were made on the basis of our hypothesis that children understand the formal and functional constraints of stroke patterns:

1. If children learning Chinese do in fact understand the formal constraints of written Chinese (as indicated by the results of Study 2), they will create significantly more pseudowords than nonwords.
2. If they understand the functions of stroke patterns, they will systematically choose the appropriate semantic radicals when creating the new words to designate the peculiar objects and will selectively choose as the second component to form a name for the objects those stroke patterns which are phonological components.
3. If they understand the function of phonological components, they will invent a pronunciation for their pseudowords which is related to the phonological component that they choose.

We also expected children's understanding of the formal and functional constraints to increase with age.

8.1.2 Subjects

Thirty-six children, aged 4 - 9 participated in this study

8.1.3 The task

The task was organised in the form of a game. A boy, Dung-dung was introduced. One day, he decided to visit a planet that no one had ever visited before. When he arrived there, he found many strange objects - pictures of which were shown to the children without any mention of a name - a spaceman, a ship, a house, a pool, an animal and some unusual flowers (singular and plural forms are not represented in nouns in Chinese). The children were encouraged to help Dung-dung designate the strange objects using the stroke-patterns provided.

Design

Twelve stroke-patterns were provided to help the children to invent new words. The stroke-patterns chosen were very common and children were expected to recognise them without difficulty. The semantic radicals in this study were unpronounceable graphic patterns and could not stand on their own as characters. The semantic radicals chosen occupy fixed positions in compound words. For example, the semantic radicals < 亻 >; < 彳 > and < 冫 > should be placed on the left; < 宀 > and < 艹 > should be placed on the top; and < 辶 > must be put at the bottom of the word. The position of this set of stroke-patterns decided the orthographic acceptability of the invented words; violating the positional rule resulted in the formation of nonwords, and therefore, could not be accepted as proper 'words'.

The function of semantic radicals was to provide meanings for the invented words. The six semantic radicals indicated the categories of human beings < 亻 >, water < 氵 >, animal < 犭 >, house < 宀 >, ship < 舟 >, and flower < 艹 >, which corresponded to the objects in the pictures - a strange person, a weird pool, an unusual animal, a strange house, an unreal boat, and some strange flowers. If children used appropriate semantic radicals, then the objects could be clearly classified into categories.

The other six stroke-patterns were treated as phonological components. They were pronounceable, and their meaning were not related to the objects of the pictures. These six stroke-patterns, which were also characters, could provide phonological clues to the invented words: *half* < 半 > [bun³], *east* < 東 > [dung¹], *old* < 古 > [gu²], *up* < 上 > [seong⁶], *very* < 太 > [tai³], *down* < 下 > [ha⁶].

In order to form a pseudoword, children would have to combine a semantic radical and a phonological component. Combining two semantic radicals would not form a compound word, because an orthographic rule would be violated: in all Chinese compound words at least one component must be a simple character. Combining the phonological components could not form a compound word either: in all Chinese compound words, one component has to be a semantic radical.

8.1.4 Materials

A cover story picture, six pictures with strange objects and twelve stroke-patterns, were used in this task:

1. A cover story picture with a young boy, Dung-dung sitting on a rocket ready to set off.
2. Six pictures with objects drawn in an idiosyncratic way - a spaceman, a house, a ship, a plant, an animal and a pool (Figure 8.1).



1. Twelve stroke-patterns were provided. They were written on transparencies so that children could combine the stroke-patterns to form words more efficiently.

Semantic Radicals	Phonological Components
< 亻 > - spaceman	< 半 > - half [bun ³]
< 氵 > - pool	< 東 > - east [dung ¹]
< 犭 > - animal	< 古 > - old [gu ²]
< 宀 > - house	< 上 > - up [seong ⁶]
< 舟 > - ship	< 太 > - very [tai ³]
< 艹 > - flower	< 下 > - down [ha ⁶]

8.1.5 Procedure

The experimenter introduced Dung-dung with the help of the cover story picture. She then displayed all the six pictures, discussed with the child what he/she could see in the pictures and then put the twelve transparencies on the table. Children were encouraged to make six pseudowords by matching two transparencies in whatever way they liked. When they finished matching the components, they copied the words into the space provided in each picture. Altogether they had to write six words, one word corresponding to each picture. They were then encouraged to pronounce the words that they had created.

The following instructions were given to each child:

“There are six pictures and twelve transparencies. Try to use two transparencies to make names for these objects.”

“You can change your mind and rearrange the transparencies again if you want.”

“Please copy the words from the transparencies to the square on the picture.”

“How would you like to pronounce this word?”

8.1.6 Results

8.1.6.1 *Understanding formal constraints: the position of semantic radicals and phonological components*

All the children put two stroke-patterns together to form their words as requested by the experimenter. However, only 77% of the responses could be accepted as pseudowords: they consisted of a semantic radical and a phonological component and the components were placed in their legal positions. The rest of the responses were classified as nonwords. The nonwords were divided into two categories: Nonwords A, where the semantic radical and the phonological component were placed in illegal positions, and Nonwords B, where two semantic radicals or two phonological components were used to form a pseudoword. Table 8.1 presents the percentage of pseudowords and nonwords in creative spelling by age group.

Table 8-1 Percentage of pseudowords and nonwords in creative spelling by age group

	Pseudowords	Nonwords (A)*	Nonwords (B)*
4-year-olds	25	25	50
5-year-olds	73	3	24
6-year-olds	83	0	17
7-year-olds	85	0	15
8-year-olds	94	2	4
9-year-olds	100	0	0
Total	76	5	19

* Nonwords (A) - semantic radicals and phonological components in illegal positions

Nonwords (B) - two semantic radicals or two phonological components

Apart from the 4-year-olds, the majority of the children formed pseudowords appropriately. Further analysis was carried out by giving children a mark if they could form a pseudoword - that is, with both the semantic radicals and phonological components being in their legal positions, and the maximum score was six. Table 8.2 presents the mean scores of pseudowords by age group. A one-way ANOVA was conducted to look at the age effect on the number of pseudowords as the dependent variable. A significant age effect was found, $F(5,30)=4.56$, $p<0.05$. Post hoc comparison by Student-Newman-Keuls test revealed that the 4-year-olds scored significantly lower than the other age groups.

Table 8-2 Mean out of six of pseudowords by age group

4-year-olds	1.5
5-year-olds	4.6
6-year-olds	5.0
7-year-olds	5.1
8-year-olds	5.6
9-year-olds	6.0

In the subsequent sections examining children's understanding of the functioning of the semantic radicals and the phonological components, only pseudowords were included, while nonwords (A) and (B) were excluded from further analysis.

8.1.6.2 Understanding functional constraints: the use of semantic radicals to represent meaning

The analysis of the creation of pseudowords only indicates that children use the correct sort of element in the correct place. However, if children understand the function of the semantic radicals, they should be able to choose not any semantic radical, but rather the appropriate radical systematically when inventing the new words. There were twelve stroke-patterns presented to children altogether. Six of them were semantic radicals

which indicated the general category of the objects in the pictures. They were $\langle \text{亻} \rangle$ for *spaceman*, $\langle \text{氵} \rangle$ for *pool*, $\langle \text{犭} \rangle$ for *animal*, $\langle \text{宀} \rangle$ for *house*, $\langle \text{舟} \rangle$ for *ship* and $\langle \text{艹} \rangle$ for *flower*. If children were aware of the function of the semantic radicals, then they might match the semantic radicals to the pictures in order to indicate meaning. Table 8.3 presents the frequency of semantic radicals correctly chosen to write the names of the peculiar objects in pseudowords. Nonwords (A) and (B) were excluded from the table. A series of χ^2 tests indicated that the probability of choice of the correct semantic radical for each particular object was significantly higher than that which would be expected at chance level.

Table 8-3 Frequency of selection of semantic radicals by picture

Pictures	Semantic Radicals						N	χ^2	df	P
	亻	氵	犭	宀	舟	艹				
spaceman	20*	3	1	4	2	0	30	41.7	4	.000
pool	2	18*	1	2	2	4	29	44.0	5	.000
animal	2	3	20*	3	2	0	30	41.0	4	.000
house	2	3	0	20*	4	2	31	38.8	4	.000
ship	3	2	2	2	16*	4	29	31.6	5	.000
flower	2	3	1	3	1	21*	31	59.0	5	.000

* correct match

Further analysis was carried out by giving one score for every correct semantic match, and the maximum score was six. Pro-rata scores were calculated, as nonwords were excluded from the analysis. Table 8.4 presents the mean and standard deviations of correct radicals by age group. A one-way ANOVA was conducted to look at the age effect with choice of correct radicals as the dependent variable. No age effect was found in matching radicals correctly to corresponding pictures, $F(5,30)=2.4$, although an age trend was apparent.

Table 8-4 Mean out of six and standard deviation of correct radicals by age group

4-year-olds (n=2)	0
5-year-olds (n=6)	2.7 (2.7)
6-year-olds (n=6)	3.3 (2.3)
7-year-olds (n=8)	3.6 (2.4)
8-year-olds (n=8)	4.0 (2.3)
9-year-olds (n=6)	5.6 (0.8)

The four-year-olds did not match any of the semantic radicals to the appropriate pictures. There was no trace of the 4-year-olds having any idea of the function of semantic radicals. The 5-, 6-, and 7-year-olds obtained approximately half of the matches correctly, and the older children, the 9-year-olds, demonstrated a particularly clear awareness of the function of semantic radicals by matching most of the semantic radicals correctly to the pictures. Post hoc examination by Student-Newman-Keuls test revealed that the 4-year-olds were significantly different from the other age groups, while the other age groups did not differ from each other significantly.

8.1.6.3 Understanding functional constraints: the use of phonological component in reading

In contrast to the semantic radicals, the phonological components could not be systematically related to the objects when the children were inventing their new words. Therefore, a random distribution is expected in the selection of phonological components. Table 8.5 presents the frequencies of children' selection of phonological components for each picture. Nonwords (A) and (B) were not included in the frequency table.

Table 8-5 Frequency of selection of phonological components by picture

Pictures	Phonological Components						<i>N</i>	χ^2	<i>df</i>	<i>p</i>
	半	東	古	上	太	下				
spaceman	6	5	4	2	11	2	30	11.2	5	.05*
pool	5	5	5	3	3	8	29	3.5	5	.63
animal	6	4	6	7	6	1	30	4.8	5	.44
house	4	7	4	3	5	8	31	3.6	5	.60
ship	9	4	6	3	3	4	29	5.6	5	.35
flower	4	6	5	8	4	4	31	2.5	5	.78

1. < 半 > half [bun³]2. < 東 > - east [dung¹]3. < 古 > - old [gu²]4. < 上 > - up [seong⁶]5. < 太 > - very [tai³]6. < 下 > - down [ha⁶]**p* = .0476, *p* < 0.05

The frequencies of selection of the different phonological components did not differ from chance level, except in one case: there was a tendency for children to choose the phonological component < 太 > [tai³] to match with the picture “spaceman”, because the first character for the word *spaceman* [太空人], is also < 太 > [tai³].

The random allocation of phonological components is of interest because it offers a strong contrast with the systematic choice of semantic radicals. However, it is only the first step. The focus of the analysis was to find out whether the children realized that these stroke-patterns have the particular function of providing clues to pronunciation. Thus children’s reading of each of their invented words was classified according to the type of strategy that they seemed to use to generate the pronunciation. Three main strategies could be identified without difficulty.

1. Strategies based on the phonological component

Children using this strategy most of the time read the new word by derivation (that is, they pronounced it in the same way that the phonological component is pronounced on its own). Other responses appeared to be intentional alterations of the pronunciation (such as changing the vowel or the tone). Finally, other readings were based on analogy (that is, the children matched the pronunciation of the new word to that of another word which shared the same phonological component). A total of 33% of the words were read on the basis of a phonological strategy.

2. Strategies based on the semantic component

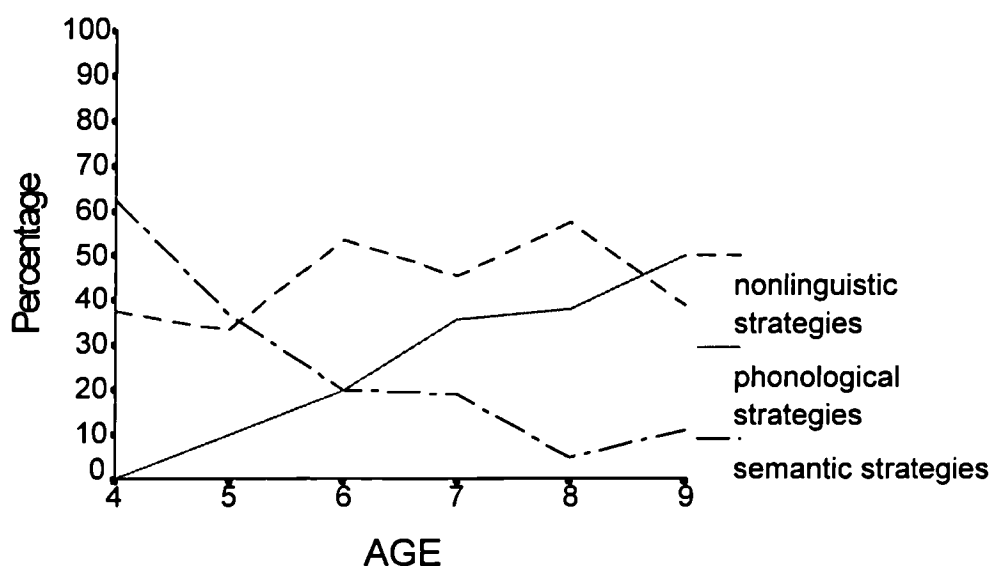
Although the semantic radicals were not pronounceable in this experiment, some children named the objects by referring to their semantic categories, for example, a flower, a ship, or they gave names of equivalent objects, for example, 'a dog' for the strange *animal*, and a river, sea, water for *pool*, and these words share the same radical components, too. A total of 18% of the pronunciation was based on semantic consideration.

3. The use of nonlinguistic clues to generate the pronunciation

Some children took into account the context of the task more than the word they had written when they were asked to read their invented words. A variety of responses was observed here. For example, some children would look for details in the pictures and use the Chinese words for those details when reading their invented words. One child used nonword sounds because the names of the objects were to be spoken in another planet. It is possible that these children did not see their task as a matter of reading what they had written; in any case, their pronunciation did not reflect the use of either the phonological component or the semantic radical in their invented words for generating the pronunciation. A total of 49% of the responses were classified in this category.

Figure 8.2 presents the percentage of children using different strategies for reading the words that they had invented by age group. There were children in all age groups who pronounced words without looking into the orthographic information they had themselves generated. There is an upward trend for the phonological strategies, indicating older children are more likely to refer to the phonological components for pronunciation. At the same time, there is a downward trend for semantic strategies. Although younger children relied heavily on the semantic information even for pronunciation, older children gave up this strategy, probably recognizing implicitly that the semantic radical can never provide any phonological information for pronunciation.

Figure 8.2 Percentage of children using different strategies in pronouncing the pseudowords by age group



Further analysis was conducted by giving the children a score on the use of a phonological strategy according to the number of words read in this way. The scores were pro-rated as nonwords were excluded from the analysis. A one-way ANOVA was conducted looking at the age effect with the number of responses using the phonological strategies as the dependent variable. The age effect was not significant $F(5,30)=2.27, p=0.07$.

8.1.7 Discussion

8.1.7.1 Understanding formal constraints: the position of semantic radicals and phonological components

Most of the children, apart from the 4-year-olds, were able to create pseudowords as names for the strange objects. They combined a semantic radical and a phonological component to form compound words. In addition, all the stroke-patterns were placed in their legal positions. A significant age trend was traced which implied that older children have a better understanding of the formal constraints of written Chinese than younger children.

A total of 24% of the responses were nonwords, and they were mainly created by the youngest children who have very limited exposure to the Chinese written language. Some five-year-olds and six-year-olds could not match the semantic radicals with the pictures, indicating that they have very little idea that some parts of a word can convey meaning, yet they were able to place most of the stroke-patterns into their appropriate place in order to form pseudowords and not nonwords. This implies that the positioning of stroke-patterns may be the first group of rules that children build up to understand the written language, and this knowledge is developed before the recognition that some stroke-patterns can convey meaning. The nine-year-olds made no mistakes at all in placing the stroke-patterns according to their correct positioning, and a ceiling effect is observed at this age.

8.1.7.2 Understanding functional constraints: the use of semantic radicals to represent meaning

Semantic radicals convey meaning and the older children were able to relate the radicals correctly to the appropriate pictures by referring to the general categories of meaning. When they were presented with the stroke-patterns, older children tended to match the semantic radicals with the appropriate pictures first, and then looked for the

phonological components. It appears that acquiring the rules for semantic radicals begins at a fairly early stage of literacy development.

8.1.7.3 Understanding functional constraints: the use of phonological components in reading

There is a random dispersion of phonological components allocated to different pictures. Unlike the semantic radicals, there is no relationship between the phonological components and the pictures. The patterned distribution of semantic radicals and the random distribution of phonological components allocated to different pictures demonstrated that children are aware of the different functions of the semantic radicals and phonological components. When the older children first looked at the stroke-patterns, they had recognised the radicals as the semantic components. They picked up the radicals, and then allocated them to the corresponding pictures. They were left with other components which were also required for word formation, something to be put into the words too, because the radicals alone could not be words by themselves. It was difficult for the children to decide which one to match with which radicals or pictures because they did not yet have a clue about how to pronounce the words. The only way to do it was by random distribution which explains why there is no pattern between the phonological components, the corresponding pictures and the semantic radicals.

When they had finished writing the words on the pictures, they were asked to pronounce them. Most of the children found it harder to make up the pronunciation than spell the words. This may imply that children have grasped the rules underlying the orthographic structure of Chinese words better than applying rules in pronunciation. The mean scores for pronouncing the words using phonological clues steadily increased with age; and phonological rules were only applied consistently in reading by the nine-year-olds.

In this study, the combination of some of the stroke-patterns resulted in real words, they were < 𠂇 > + < 古 > = < 苦 > and < 𠂇 > + < 古 > = < 估 >. Some older children were aware of these combinations and formed the real words, and found it more difficult to match with the pictures because the meaning of the real words did not match with the objects in the pictures. This confounding factor would be modified in this main study by replacing some stroke-patterns so that no combination could result in real words.

8.2 The Main Study

There were altogether two tasks in the main study. The first task was a replication of the pilot study, and the second task was a newly designed experiment to investigate the relationship between the positioning and the functioning of stroke-patterns. Task 1 was always conducted first followed by Task 2.

8.2.1 Task 1

In this replication study, some new stroke-patterns were introduced to avoid the confounding factor of forming real words in any combinations. Children were required to invent five words instead of six in this study.

8.2.1.1 Rationale

Refer to Section 8.1.1 for a full description of the rationale of the study. Although age trends could be traced in the utilization of semantic radicals for meaning and phonological components for pronunciation, the age effect was not significant in the pilot study. It was suspected that the insignificant results were caused by the small sample. In this study, the number of subjects was increased in each age group, and it was predicted that significant age trends could be traced in utilizing the phonological components for pronunciation and semantic radicals for meaning.

8.2.1.2 Subjects

One hundred children, age 5 - 9 participated in this study (for a full description, see Section 5.3.2 in Chapter 5).

8.2.1.3 Design

Refer to Section 8.1.3 for a full description of the design.

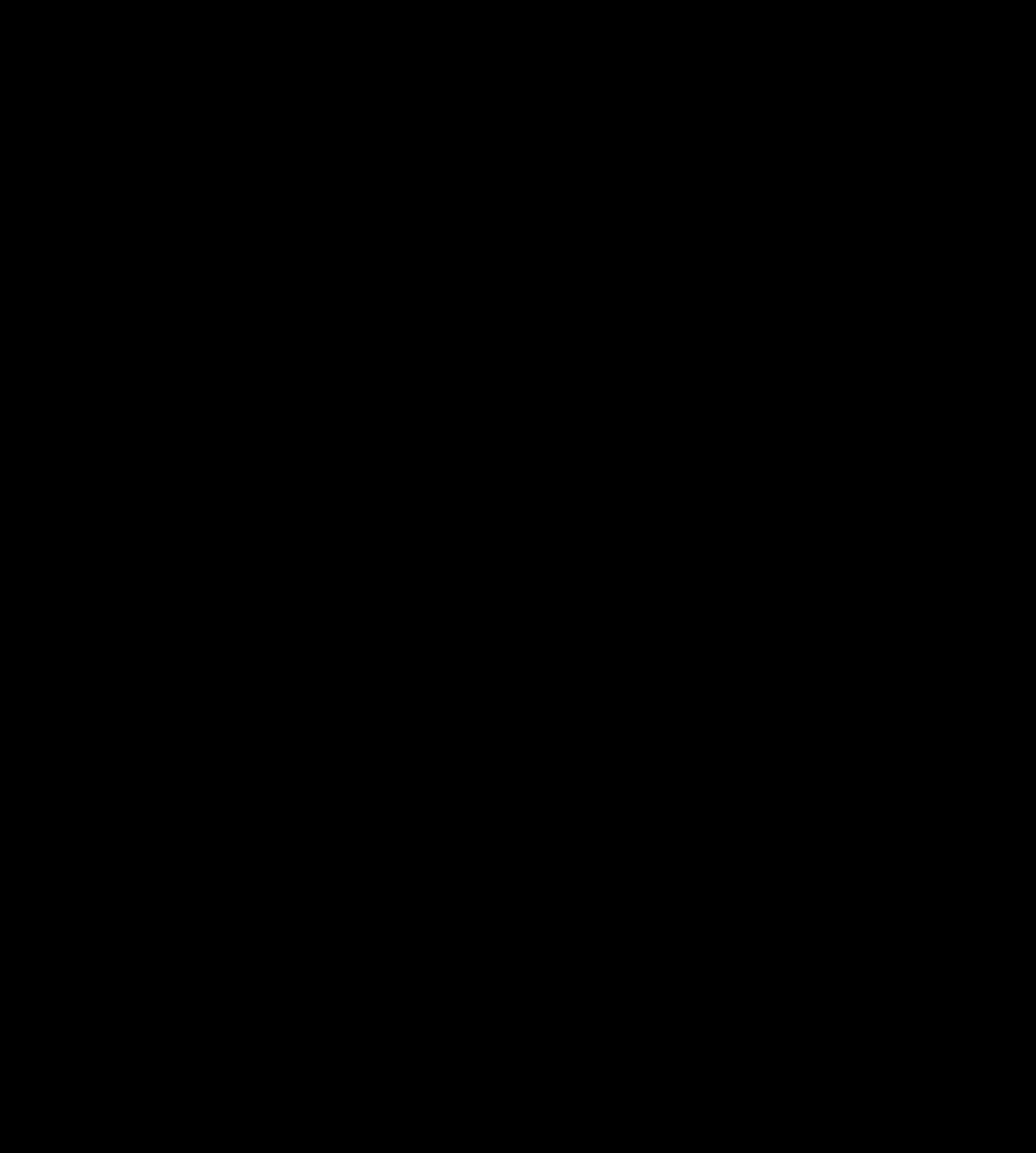
8.2.1.4 Materials

A cover story picture, five pictures with strange objects and ten stroke-patterns were used in this task:

1. A cover story picture with a young boy, Dung-dung sitting on a rocket ready to set off.
2. Five pictures with objects drawn in an idiosyncratic way - a spaceman, a house, a plant, an animal and a pool (Figure 8.3).
3. Ten stroke-patterns were provided. They were written on transparencies so that children could combine the stroke-patterns to form words more efficiently.

Semantic Radicals	Phonological Components
< 亻 > - spaceman	< 不 > - no [bud ⁷]
< 氵 > - pool	< 有 > - have [yau ⁵]
< 犭 > - animal	< 小 > - little [siu ²]
< 宀 > - house	< 上 > - up [seong ⁶]
< 艹 > - flower	< 年 > - year [ning ⁴]

IMAGE REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



8.2.1.5 Procedure

Refer to Section 8.1.5 for a full description of the procedure. Instead of forcing children to combine two stroke-patterns, children were free to use any number of stroke-patterns. The following instructions were given:

“Dung-dung found some strange objects, do you know what they are?” The objects, though strange, can be easily recognized by the children and experimenter does not give any clue in the names of the objects.

“Dung-dung wants to give them names. Here are the stroke-patterns that Dung-dung knows, and he wants to use them in order to make names for these objects that can only be found in this planet. Can you help him to make the names?”

“A spaceman comes and asks Dung-dung how to pronounce these words. Can you tell Dung-dung how to pronounce these words?”

The experimenter noted the invented words and then children were asked to pronounce the words one by one.

8.2.1.6 Results

8.2.1.6.1 Understanding formal constraints and the position of semantic radicals and phonological components

Table 8.6 presents the percentage of pseudowords and nonwords in the creative spelling by age group. Since the children were free to choose the number of stroke-patterns in forming the new words, the responses consisted of one stroke-pattern (1.6%), two stroke-patterns (98%), or three stroke-patterns (0.4%). A total of 4% of the responses failed to put the stroke-patterns in their legal positions (Nonwords A), and 14% of the responses consisted of either two semantic radicals or two phonological components (Nonwords B).

Table 8-6 Percentage of pseudowords and nonwords in creative spelling by age group

	Pseudowords	Nonwords (A)*	Nonwords (B)*
5-year-olds	49	12	29
6-year-olds	82	6	12
7-year-olds	92	1	7
8-year-olds	84	0	16
9-year-olds	95	1	4
Total	80	4	14

* Nonwords (A) - semantic radicals and phonological components in illegal positions
 Nonwords (B) - two semantic radicals or two phonological components

Further analysis was carried out by considering the frequency of pseudoword creation by age group. Children were given one mark if they could form a pseudoword, and the maximum score was five. Table 8.7 presents the mean and standard deviation of pseudowords. A one-way ANOVA was conducted to look at the age effect with the number of pseudowords as the dependent variable. A significant age effect was found, $F(4,99)=8.11$, $p<0.001$. Post hoc comparison by Scheffe test revealed that the 5-year-olds scored significantly lower than the other age groups.

Table 8-7 Mean out of five and standard deviation of pseudowords

5-year-olds	1.4 (1.0)
6-year-olds	3.8 (1.3)
7-year-olds	4.1 (1.2)
8-year-olds	4.0 (1.4)
9-year-olds	4.5 (1.1)

In the subsequent sections examining children's understanding of the functioning of the semantic radicals and the phonological components, only pseudowords formed by combining two stroke-patterns were included in further analysis.

8.2.1.6.2 Understanding functional constraints: the use of semantic radicals to represent meaning

Table 8.8 presents the frequency for semantic radicals correctly chosen to write the names of the peculiar objects in pseudowords. The frequency of correct matches accounted for 70% of the responses. A series of χ^2 tests indicated that the probability of choice of the correct semantic radical for each particular object was significantly higher than would be expected by chance. The other matches excluding the correct match were randomly distributed which showed that there was no misleading semantic meaning provided by other semantic radicals.

Table 8-8 Frequency of selection of semantic radicals by picture

Pictures	Semantic Radicals					<i>N</i>	χ^2	<i>df</i>	<i>P</i>
	𠂇	亅	𠂇	𠂇	𠂇				
flower	53*	6	5	9	4	77	115.7	4	.000
spaceman	7	57*	9	4	3	80	132.8	4	.000
house	7	4	64*	4	8	87	156.7	4	.000
animal	9	6	3	53*	9	80	108.5	4	.000
pool	3	8	9	5	63*	88	147.7	4	.000

* correct match

Further analysis was carried out to consider the age effect on the correct selection of semantic radical. One score was given for every correct semantic match and the maximum score was five. Pro-rata scores were calculated, as nonwords were excluded from the analysis. Table 8.9 presents the mean scores and standard deviations of correct selection semantic radicals by age group. A one-way ANOVA was conducted to look at the age effect with choice of correct radicals as the dependent variable. A significant age effect was found in matching radicals correctly to corresponding pictures, $F(4,99)=17.67$, $p<0.001$. Post hoc examination by Scheffe test demonstrated that the 5-year-olds scored significantly lower than all the other age groups, and that six-year-olds scored significantly lower than the eight-year-olds.

Table 8-9 Mean out of five and standard deviation of correct selection of semantic radicals by age group

5-year-olds	1.0 (1.4)
6-year-olds	2.8 (1.8)
7-year-olds	4.1 (1.6)
8-year-olds	4.3 (1.3)
9-year-olds	4.3 (2.0)

8.2.1.6.3 Understanding functional constraints: the use of the phonological components in reading

A random distribution is expected in the selection of phonological components to match with the pictures. Table 8.10 presents the frequencies of children's selection of phonological components for each picture.

Table 8-10 Frequency of selection of phonological components by picture

Pictures	Phonological Components					<i>N</i>	<i>X</i> ²	<i>df</i>	<i>p</i>
	年	有	不	小	上				
flower	20	11	18	14	13	76	3.6	4	0.46
spaceman	23	12	16	14	15	80	4.4	4	0.36
house	9	23	24	13	18	87	9.5	4	0.05*
animal	21	18	14	14	13	80	2.9	4	0.58
pool	9	19	14	21	25	88	8.8	4	0.07

1. < 年 > - year [ning⁴]

2. < 有 > - have [yau⁵]

3. < 不 > - no [bud⁷]

4. < 小 > - little [siu²]

5. < 上 > - up [seong⁶]

**p* = .049, *p* < 0.05

Similar to the finding in the pilot study, no clear pattern emerged in matching any particular phonological components to different pictures. The frequencies of selection of the different phonological components did not differ from chance level, except for a

tendency for children to reject the phonological component < 年 > - year [ning⁴] to match with the picture “house” and “pool”.

There were three main strategies children adopted when they were asked to pronounce the words that they had created: 1) Some of the children referred to the phonological components to pronounce the words (37%); 2) some of them referred to the semantic radicals to pronounce the words (26%); and 3) some of them paid attention to the situation, rather than to the stroke-patterns in pronunciation (37%). Refer to Section 8.1.6.3 for a full description of the strategies adopted by children in pronunciation.

Figure 8.4 Percentage of children using different strategies in pronouncing the pseudowords by age group

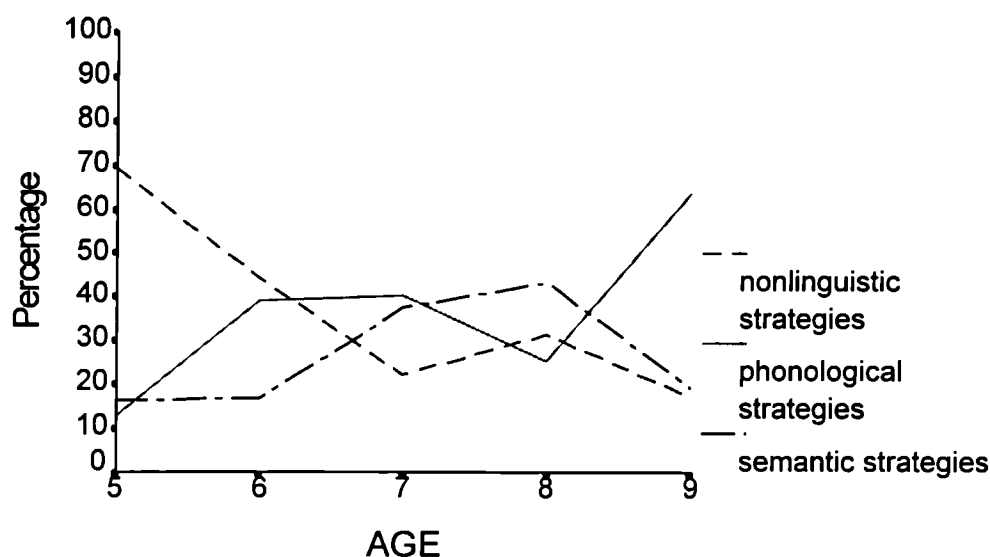


Figure 8.4 presents the percentage of children using different strategies in pronouncing the pseudowords by age group. Different age groups relied on different strategies for pronunciations. Five- and six-year-olds preferred to pronounce words by non-linguistic strategies. Seven- and eight- year-olds used all three strategies equally. Nine-year-olds began to use the phonological strategies predominantly.

Further analysis was conducted by giving the children a score on the use of the phonological strategy according to the number of words read in this way, the maximum score being five. The scores were pro-rated, as nonwords were excluded from the analysis. A one-way ANOVA was conducted looking at the age effect with the number of responses using the phonological strategies as the dependent variable. A significant age effect was found, $F(4,97)=5.54$, $p<0.001$. There is an upward trend for the phonological strategies, indicating older children are more likely to refer to the phonological components for pronunciation. Post hoc examination by Scheffe test revealed that the 5 and 8-year-olds scored significantly lower than the 9-year-olds.

8.2.1.7 Discussion

The results in this task lent support to the findings in the pilot study, as comparable results were obtained in both studies. Children aged five to nine were able to form more than 80% of the pseudowords. They were aware of the fixed positions of the semantic radicals and used this knowledge to help them to form pseudowords. A significant age trend was found in the ability to form pseudowords instead of nonwords in both studies.

Children aged five to nine were able to match about 70% of the semantic radicals correctly to the corresponding pictures. The distribution of the semantic radicals was significantly different from chance, and a significant age trend was traced in the main study, demonstrating that older children were more able to utilize the semantic radicals to represent meaning.

Children used different strategies in pronouncing the pseudowords in both studies. Younger children used non-linguistic and semantic strategies in trying to pronounce the pseudowords, and only the nine-year-olds consistently referred to the phonological

components for pronunciation. In Chinese, the relationship between the phonological components and pronunciation is rather opaque. According to Zhu's (1987) statistical analysis, the predictive accuracy of the pronunciation of an ideophonetic compound from its phonetic is about 44%. However, the predictive accuracy drops to 19% if frequency of usage is taken into consideration. In other words, ideophonetic compounds are more 'phonologically regular' in low frequency words than in high frequency words. Younger children may have been exposed to high frequency words before having acquired the Orthographic-Phonology correspondence rules in pronouncing Chinese words. An age trend was traced in the main study, showing that older children were more capable of utilizing the phonological components for pronunciation. In future studies, we recommend that the utilization of phonological components in pronunciation by older age groups, such as ten, or eleven years old children should be explored further because older children are more likely to have been exposed to more low frequency words, and they may then acquire the phonological skills similar to those of adult Chinese readers.

Furthermore, only weak evidence was gathered that children utilize the phonological components in pronunciation. It is not clear if children are aware of the fact that the phonological component, that is the non-radical component, is the only unit that provides phonological information since they were also the only pronounceable units in this study. The performance of the children may be achieved with the help of the unpronounceable semantic radicals. If we look from a different perspective, it is the insistence of children to disregard the phonological components for pronunciation, even though they are the only pronounceable units, that demands further investigation. The phonological components in this study were simple characters and all the children knew how to pronounce them. The reasons why they did not utilize the phonological information in the phonological components may be due to the following reasons: they may not have known that phonological information can be elicited in the words; they may believe that the phonological information can be found in the semantic radicals because about 26% of the responses in the main study used the semantic radicals for pronunciation even though they were not pronounceable; and it is also possible that they

were not forming ideophonic compound words, but ideocompound words where both components were intended to represent meaning and not pronunciation.

To resolve all these queries, an additional task was included in the main study. In the following task, all the stroke-patterns were simple characters which can be pronounced and at the same time have meaning. Children were free to choose any stroke-patterns to form different combinations. Since all stroke-patterns are pronounceable, it may provide a clearer picture as to which components children would use for pronunciation.

8.2.2 Task 2

8.2.2.1 The relationship between the positioning and functioning of stroke-patterns

In task 2, a similar investigation into the functions of stroke-patterns was carried out. It was stated at the beginning of this chapter that the functions of stroke-patterns are indicated by their position within a character. In task 1, evidence was gathered that children are aware of the semantic and phonological functions of the stroke-patterns. However, it is still unclear if children can decide the functions of the stroke-patterns purely on position, irrespective of the stroke-patterns.

There are some Chinese stroke-patterns that can appear in almost every position in a character. However, if they are to function as semantic radicals, then they have a fixed position. For example, the stroke-pattern *woman* < 女 > can appear on the left, right, and bottom of a character, but as a semantic radical, it can only appear in the left on a left-right combination, or at the bottom in a top-bottom combination. The relation between the position and the function of stroke-patterns is so closely related that some psychologists regard the radical on the right as the phonological component, and the

radical on the left to be the semantic radical (Seidenberg, 1985; Flores d'Arcais, 1992, 1995). Again, this explanation simplifies the situation because there are exceptions in the relation between the positions and the functions. However, knowing the basic semantic radical position rule, such as that most of the semantic radicals can be found on the left in a left-right combination, helps children to decide the different representational functions of the stroke-patterns which can appear in different positions.

In task 1, there was some indication that children might confuse the function of the semantic radicals. Although most of them understood that the semantic radicals offered clues to meaning, they were not aware that semantic radicals are never pronounced and are unlikely to provide any kind of phonological information in a word. About 25% of the invented words were read by referring to their semantic categories in Task 1, although the semantic radicals were unpronounceable. In this task, pronounceable semantic radicals were used in order to allow a better investigation.

The two predictions for Task 2 were as follows:

1. If children understand that the position of the stroke-patterns is an indication of their representational functions, they will allocate the semantic radicals and the phonological components to their respective positions.
2. If they understand the function of phonological components, they will invent a pronunciation for their pseudowords which is related to the phonological component that they choose.

8.2.2.2 Subjects

The subjects in this task were the same as those in Task 1.

8.2.2.3 Design

Children were presented with another set of pictures and stroke-patterns. Five pictures of peculiarly looking but clearly recognizable objects - a car, an insect, a woman, a fish and a bird were as used as those objects for which the children were asked to invent names.

Ten stroke-patterns were provided. All of them were simple characters which could be pronounced, and conveyed meaning on their own. When they are used as stroke-patterns in compound characters, they can be found in different positions and serve as either semantic radicals or phonological components. Therefore, no combinations of the stroke-patterns could result in nonwords in this task.

Although this task placed no orthographic constraints on the process of inventing pseudowords, a deeper level of understanding of the relationship between semantic radical and its position was necessary in order to create a meaningful word. The same stroke-pattern can have either the function of semantic radical or that of phonological component, depending solely on its positioning in the word. Children could no longer depend on the graphic structure of the semantic radical to help them to decide the position, as they possibly did in Task 1. In representing meaning, children have to choose the correct semantic radical and then place them in the correct semantic radical positions.

The meaning of five semantic radicals, < 車 > - car [tse¹], < 鳥 > - bird [niu⁵], < 虫 > - insect [tsung⁴], < 魚 > - fish [jy⁴], < 女 > - woman [noey⁵] was related to the meaning of the pictures. If the correct semantic radicals were used in inventing words, they represented the general category the objects belonged to.

In pronouncing the pseudowords, children have to avoid the pronunciation of the semantic radicals, and to take note of the phonological components. This knowledge, however, is not explicit even to adult readers

8.2.2.4 Materials

Five pictures with strange objects and ten stroke-patterns were used in this task:

1. Five pictures with objects drawn in an idiosyncratic way – the strange objects were a huge woman, a bird with two heads, a boot-shaped car, an insect and a fish with many eyes (Figure 8.5).
2. Ten stroke-patterns were provided. They were written on transparencies so that children could combine the stroke-patterns to form words more efficiently.

Semantic radicals	Phonological components
< 車 > - car [tse ¹]	< 米 > - rice [mai ⁵]
< 鳥 > - bird [niu ⁵]	< 刀 > - knife [dou ¹]
< 女 > - woman [noey ⁵]	< 用 > - use [yung ⁶]
< 虫 > - insect [tsung ⁴]	< 元 > - beginning [jyn ⁴]
< 魚 > - fish [jy ⁴]	< 天 > - sky [tin ¹]

IMAGE REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES

8.2.2.5 Procedure

The same procedure in Task 1 was adopted in Task 2.

8.2.2.6 Results

Very few children used one stroke pattern, or combined three stroke-patterns together (2%). A total of 98% of the invented pseudowords consisted of two stroke-patterns using left-right combinations. However, some pseudowords were made by combining either two semantic radicals or two phonological components (22%); these data were impossible to analyse unless children pointed out which stroke-patterns they regarded as the semantic radicals. Thus the following analysis was based on the remaining 76% of the responses where children combined a semantic radical and a phonological component in their creative spelling.

8.2.2.6.1 Understanding the formal and functional constraints: the use of semantic radicals to represent meaning

The five semantic radicals which indicated the general category of the objects were *car* < 車 >, *bird* < 鳥 >, *insect* < 虫 >, *fish* < 魚 >, and *woman* < 女 >. Table 8.11 presents the frequencies for semantic radicals matching the corresponding pictures. Most of the children were able to match the semantic radicals correctly to the corresponding pictures (88%). A series of χ^2 indicated the probability of choice of the correct semantic radical for each object was significantly different from chance.

Table 8-11 Frequency table for semantic radicals matching the corresponding pictures

Pictures	Semantic Radicals					N	χ^2	df	P
	車	鳥	虫	魚	女				
car	77*	1	1	6	4	89	247.12	4	.000
bird	3	61*	5	0	2	71	140.77	3	.000
insect	1	7	56*	3	0	67	123.77	3	.000
fish	3	0	3	68*	2	76	168.52	3	.000
woman	2	1	2	0	71*	76	189.79	3	.000

* correct match

The function of the semantic radicals depends on their positioning in the words. All the semantic radicals, except for < 鳥 > - bird [niu⁵] were to be placed on the left, which is also the conventional semantic radical position. If children followed the positional rule for semantic radicals, then they would place all the semantic radicals on the left. Table 8.12 presents the frequency of locating the correct semantic radicals in different positions. Most of the children put the semantic radicals on the left of the pseudowords (70%). Although the correct semantic radical position for *bird* < 鳥 > is on the right, many children simply followed the semantic radical positional rule and placed it in a wrong place .

Table 8-12 Frequency table of correct semantic radicals allocated to different positions

semantic radicals	left	right	N	χ^2	df	P
< 車 > - car	63*	14	77	28.60	1	.000
< 鳥 > - bird	35	26*	61	1.33	1	.25
< 虫 > - insect	47*	9	56	25.79	1	.000
< 魚 > - fish	49*	19	68	13.24	1	.000
< 女 > - woman	59*	12	71	31.11	1	.000

* correct match

Although 88% of the responses consisted of a correct semantic radical, the percentage dropped to 55% when the correct positioning of these semantic radicals was taken into consideration. Further analysis was conducted by giving a score to the correct semantic radicals placed in the correct location and the maximum score was five. The scores were pro-rated, as some of the pseudowords did not consist of a semantic radical and a phonological component. Table 8.13 presents the means scores and standard deviation for correct positioning of semantic radicals. A one-way ANOVA was conducted to look at the age effect with the number of correct semantic radicals as the dependent variable. A significant age effect was found, $F(4,98)=11.8$, $p<0.001$. Post hoc examination by Scheffe test showed that 5-year-olds scored significantly lower than the 7-, 8- and 9-year-olds, and 6-year-olds scored significantly lower than the 9-year-olds in matching the correct semantic radicals to the appropriate positions.

Table 8-13 Mean and standard deviation of correct positioning of semantic radicals

5-year-olds	1.6 (1.4)
6-year-olds	2.3 (1.4)
7-year-olds	2.9 (1.3)
8-year-olds	3.5 (0.8)
9-year-olds	3.9 (0.9)

8.2.2.6.2 Understanding functional constraints: the use of the phonological components in reading

The following stroke-patterns were used as phonological components, *rice* < 米 >, *knife* < 刀 >, *use* < 用 >, *beginning* < 元 > and *sky* < 天 >. If children regarded these stroke-patterns as phonological components, then they would distribute them at random as they bore no relationship to the objects in the pictures. Table 8.13 presents the frequency for phonological components matching corresponding pictures. A series of χ^2 indicated that the distribution of the phonological components was significantly different from chance in the pictures, car, bird and fish. There was a tendency to match the stroke-pattern *use* < 用 > to the picture of the car; *sky* < 天 > to the picture of the bird; and *knife* < 刀 > to the picture of the fish.

Table 8-14 Frequency table for phonological components matching the corresponding pictures

Pictures	Phonological Components					N	χ^2	df	P
	米	刀	用	元	天				
car	13	12	30	26	8	89	20.72	4	0.00*
bird	7	12	7	8	37	71	46.96	4	0.00*
insect	15	19	13	12	8	67	4.87	4	0.30
fish	20	23	9	16	8	76	11.50	4	0.02*
woman	19	9	15	19	14	76	4.53	4	0.34

1. < 米 > - rice [mai⁵] 2. < 刀 > - knife [dou¹] 3. < 用 > - use [yung⁶]

4. < 元 > - beginning [jyn⁴] 5. < 天 > - sky [tin¹]

* p < 0.05

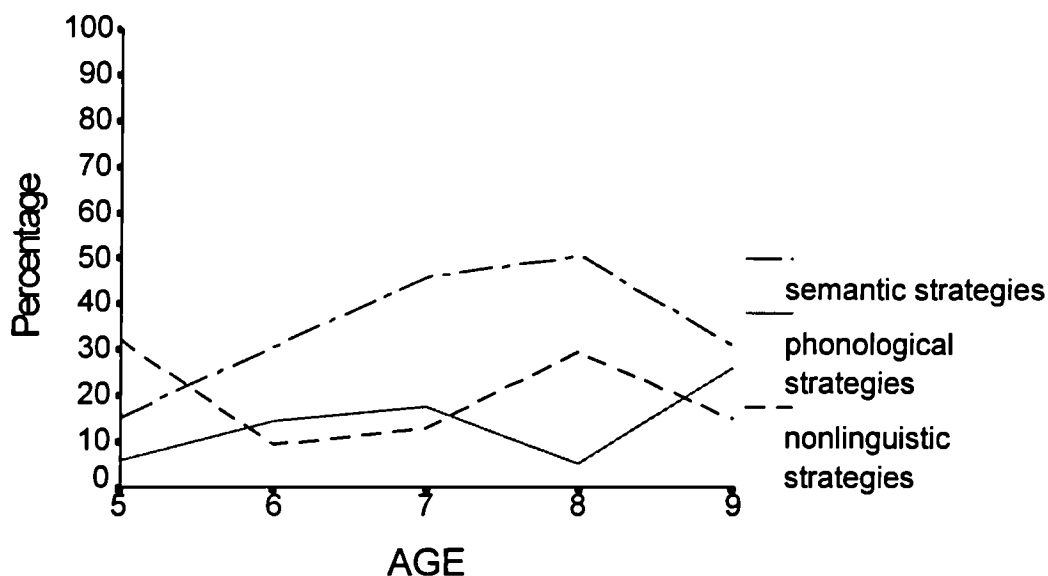
There were some children who did not regard the stroke-patterns as phonological components. The stroke-pattern *use* < 用 > was matched with the stroke-pattern *car* to mean ‘using the car’, the stroke-pattern *sky* < 天 > matched with the stroke-pattern *bird* to imply ‘a flying bird’, and the stroke-pattern *knife* < 刀 > was used with the stroke-pattern *fish* to describe the shape of the mouth of the strange fish. They formed ideocompound words where both stroke-patterns were used to convey meaning.

The strategies children adopted in pronouncing the pseudowords became very complicated. First, some children combined the two stroke-patterns to form pseudowords, but when they were asked to pronounce them, they pronounced the stroke-patterns instead of the pseudowords. These responses (7% of the pseudowords) were excluded from further analysis as they were not pronouncing the pseudowords. Second, some children in this task formed ideocompound pseudowords, instead of ideophonic pseudowords, which meant that the phonological components would not be used for pronunciation.

Although children adopted the same kind of strategies in pronouncing the pseudowords, the preference for strategies was quite different from those in Task 1. A total of 37% of the responses used the semantic radicals for pronunciation, 16% referred to the phonological components for pronunciation, and another 16% used nonlinguistic clues to generate the pronunciation. Figure 8.6 presents the percentage of children using different strategies in pronouncing the pseudowords by age group. There was, however, a decline in the use of semantic and nonlinguistic strategies, and an increase in the use of phonological components for pronunciation in the nine-year-olds.

Further analysis was carried out by giving a score to responses using the phonological clues for pronunciation and the maximum score was five. The scores were pro-rated, as the responses which did not combine a semantic radical and a phonological component, and those which pronounced the stroke-patterns instead of the pseudowords were excluded from the analysis. A one-way analysis was conducted looking at the age effect with the number of responses using the phonological strategies as the dependent variables. A significant age effect was found, $F(4,99)=2.41$, $p<0.05$. Post hoc examination by Scheffe test revealed that the 9-year-olds scored significantly higher than the 5-year-olds.

Figure 8.6 Percentage of children using different strategies in pronouncing the pseudowords by age group



8.2.2.7 Conclusion

This task was set up to investigate two aspects of children's understanding of the functional characteristics of the stroke-patterns which could not be explained satisfactorily in Task 1. The first was to examine whether children could allocate semantic radicals correctly without the assistance of graphic clues, and the second was to investigate further the reasons for not using the phonological components for pronunciation.

The results from this task provided answers to both questions. Children have acquired a deeper understanding of the relationship between the semantic radical and its positioning. They realise that most of the semantic radicals can be found on the left of a character and they use this knowledge to help them to place the stroke-patterns which do not have fixed graphic positioning. Children placed most of the semantic radicals on the left and the allocation was significantly different from chance, except for the semantic radical *bird* which was placed in an unconventional position. However, this knowledge

is learned much later than the orthographic rule according to which children can rely on the graphic structure to allocate the semantic radicals. While the 7-year-olds could place 80% of the familiar semantic radicals correctly in Task 1, only the 9-year-olds could place the same percentage of the semantic radicals correctly in Task 2. The difficulty of mastering the semantic radicals positional rule probably lies in the presence of a number of exceptions where the semantic radicals are not placed in the conventional positions.

In the enquiry into the strategies for pronunciation, it was found that most of the children are not aware that semantic radicals cannot be pronounced. It is not a surprising finding though, as many fluent Chinese readers cannot spell out this rule explicitly. However, adult Chinese readers tend to avoid pronouncing semantic radicals, while children have no such intention. In fact, many children use the semantic radicals to pronounce the ideocompound pseudowords that they have formed. Although there is no phonological rule in pronouncing the ideocompound word, its semantic radical can never provide clues for pronunciation. Therefore, almost half of the pseudowords were pronounced incorrectly in this task. Only the 9-year-olds demonstrated a tendency to reject the semantic strategies and non-linguistic strategies, and to refer to the phonological components for pronunciation. This result is consistent with the finding in the Task 1 and the pilot study.

Another possible reason for younger children using the semantic radicals for pronunciation may be caused by the context effect. Children were presented with pictures of different objects, and they might have been naming the objects instead of pronouncing the pseudowords. Therefore, the best way to examine children's pronunciation strategies is to ask them to pronounce some pseudowords in a decontextualised situation, and this setting was adopted in the following study.

There is a confounding factor in the design of this task. In general, there is a change in the size of the characters when they are used as stroke-patterns. The stroke-patterns become longer and slimmer in a left-right combination in order to allow another stroke-

pattern to fit into the square-shape which characterizes Chinese characters. However, this effect could not be produced in this task, as all the characters were written in a square shape in order that children could not get any orthographic clue as to where to put them. As a result, some younger children failed to treat the pseudowords as a compound character and pronounced the stroke-patterns separately.

To conclude, children systematically refer to semantic radicals for meaning and phonological components for pronunciation. The ability to use the semantic and phonological information in the stroke-patterns increases with age, indicating that intra-character knowledge may be crucial in the process of learning to spell.

Chapter 9

Study 4: Reading Chinese pseudowords

9.1 *The Pilot Study*

9.1.1 The phonological and semantic constraints in reading

In this study, the ability of children to identify semantic and phonological information in order to read pseudowords was examined. Evidence was gathered in the creative writing task that children were able to utilize semantic radicals to represent meaning and to use different strategies to read pseudowords that they had invented. But since children were first presented with the stroke-patterns and then made their own words, there is a possibility that the functional aspects of the stroke-patterns have been highlighted. In this study, children were to read whole words instead of fragments, and they were required to look for the linguistic components - the semantic radicals and the phonological components - which contain relevant representational functions in different words.

Some evidence has been gathered that Chinese children are aware of the relationship between speech and print. Children at the age of five can apply the one-to-one correspondence rule in reading (Chan, 1990). They count the number of syllables in the spoken words and match them with the number of characters in the written words.

In order to investigate children's understanding of the relationship between speech and print at an intra-character level, Ho (1994) gave children lists of ideophonic compound words and pseudowords to read, and found that Chinese five-year-olds already possess some rudimentary knowledge of the Orthographic-Phonology

correspondence rules, and first graders (6-year-olds) and second graders (7-year-olds) can apply their knowledge of the OPC rules in reading. However, it is still unclear if children who could pronounce the words correctly relied on a positional rule, because most of the phonological components were found on the right, or whether they were aware of the functions of the phonological components.

The results from the creative writing task in Study 3 provide evidence that the knowledge of the relationship between speech and reading develops through the following stages. Chinese children may not know that phonological information is embedded in the character in the beginning stage of learning to read, and they read words by nonlinguistic clues. This may be followed by an attempt to seek phonological information within a character, but children become confused by the representational functions of the semantic radicals. With an increase in the understanding of the Chinese writing system, older children would be able to use the clues provided by the phonological components for pronunciation. The present study investigated whether the same developmental stages could be found in reading tasks.

In Study 3, children also demonstrated an understanding of the functions of semantic radicals. The 7-year-olds matched the appropriate semantic radicals with the corresponding pictures, so that the category of the objects was appropriately represented by the semantic radicals. When they were nine, they understood that the relation between representational functions and semantic radical position when they were faced with stroke-patterns that could be used both as semantic radicals and phonological components. The present study examines if the same findings can be found in reading tasks where whole words were presented instead of fragments of stroke-patterns.

Furthermore, it is still unclear if children would regard the semantic radical as the only representational unit for meaning in compound words. They may have thought that meaning can also be conveyed by pictorial depiction of the objects, because some simple Chinese words are made in this way. In ancient times, Chinese writing was mainly ideographic in nature and symbols represented pictures of things.

Ferreiro and Teberosky (1982) suggest that children first believe that written language is a particular way of representing objects, e.g. size of words represent the height or age of different people, longer words represent larger objects, etc. They point out that although young children can differentiate drawing and writing, their interpretation of the written language is still powerfully influenced by drawing. Children find it hard to break free from the constraints that drawing imposes. They fail to conceive of an entirely different system to represent events and objects on paper. There is evidence that Chinese children are confused with the function of print and picture. Chan (1990) found that Chinese preschool children tend to 'draw' the words when they do not know how to write them. For example, some children drew a conventional house with roof, door and windows when they were asked to write the word *house*. Some children insisted on keeping the distinctive feature of an elephant - its long nose - in their creative writing.

In Study 2 where children were asked to discriminate pseudowords from nonwords, five-year-olds have already demonstrated that they would reject nonwords with illegal graphic patterns. However, it is still not known whether children, if were presented with illegal graphic patterns which represented pictorial depiction of the objects, could reject the graphic similarity and refer solely to the semantic radicals for meaning. If children are able to do so, then it implies that children understand that most of the Chinese words are no longer ideographs.

The pilot study sought to answer the following two questions: (1) Can children transfer the phonological information of a real word to read a pseudoword? (2) Can children

transfer the semantic information of a real word to understand the meaning of a pseudoword? If children can locate the phonological component in a compound word, then they can identify a pseudoword with the same phonological component which is more likely to have similar pronunciation. In the same way, if children understand that semantic information can only be found in the semantic radical, then they should be able to match a suitable pseudoword with a real word with an identical semantic radical which is more likely to have a similar meaning.

9.1.2 Subjects

Thirty-six children, aged 4 - 9 participated in this study.

9.1.3 Design

There were altogether two sessions in this experiment. In the first session, children were asked to look for a pseudoword that had the same pronunciation as the prototype. In the second session, children were asked to look for a pseudoword that had the same meaning as the prototype. The same set of material was used for both sessions.

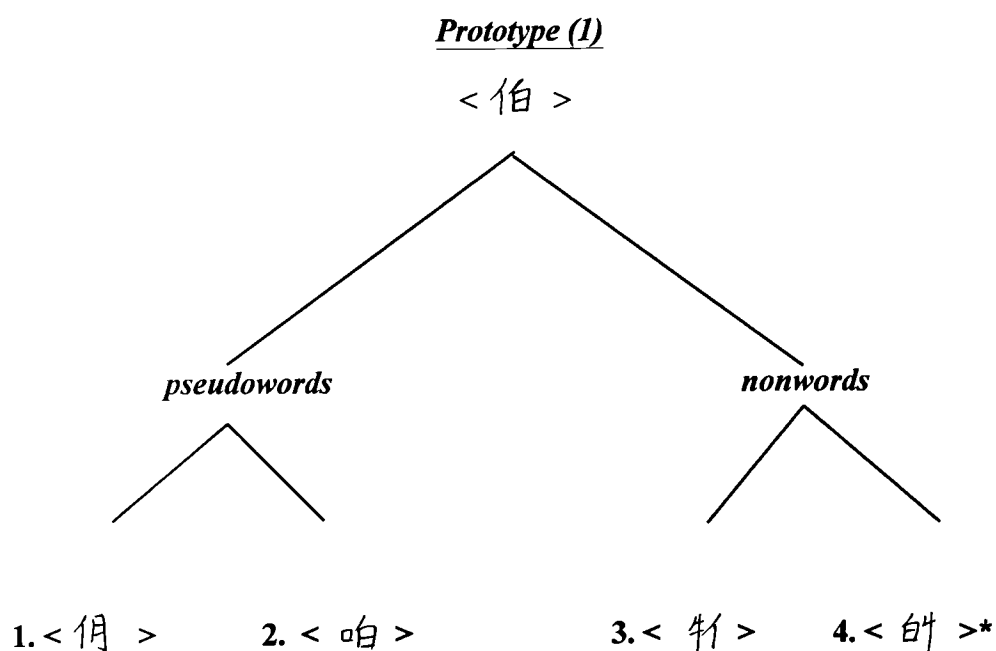
Ten real words were chosen as prototypes. For each prototype, four pseudowords/nonwords were designed as choices. The prototypes and the choices were matched in the number of strokes and the number of stroke-patterns. Two of them consisted of the same semantic radical and two of them consisted of the same phonological component as the prototype. The semantic radical and the phonological component were placed in different positions so that two choices were pseudowords and two became nonwords.

The task was designed in this way because it was assumed that Chinese children were very reluctant to read words they had not learned before. Thus the prototype, which

was a real word, was presented first, and children were only required to point to one of the choices that might provide the similar pronunciation.

The four types of choices were (1) pseudowords with matching phonological components; (2) pseudowords with matching semantic radicals; (3) nonwords with matching phonological components and (4) nonwords with matching semantic radicals (Figure 9.1).

Figure 9.1 Prototype and the four choices



* The words were presented to children in black only

9.1.4 Materials

There are altogether ten trials. For each trial, the child was presented with five cards, with a word written on each card (13 cm x 7.5 cm). The prototypes were written on red cards and the choices were written on blue cards.

9.1.5 Procedure

Session one - the phonological match

The prototypes were shown one by one. The children pronounced and explained the meaning of the prototypes. If the children did not know the pronunciation and/or the meaning of the prototypes, they were explained to them, to make sure that they knew how to say the word and to understand its meaning before the task began.

Four choices were shown and children were required to point to the word that sounded the same as the prototype. The experimenter explained to the children that the words in the choices were very difficult and that they would have to try to find the right answer but were not likely to know the words beforehand.

The following instructions were given:

“Can you tell me how to read this word (the prototype)?

“Tell me the meaning of the word.”

“I am going to show you four more words. They are very difficult words but one word is pronounced the same as the prototype. Try to choose which is the one.”

Session two - the semantic match

In session two, children were asked to find a word which meant the same as the target word. The procedure was the same as in session 1. The instruction for this task was, “Show me a word which has the same meaning as the prototype.”

9.1.6 Results

9.1.6.1 *The phonological match*

Table 9.1 presents the percentage of children's responses to the four choices in the phonological match. The majority of the responses (68%) fell into the correct choice "pseudowords with matching phonological components".

Table 9-1 Percentage of children's responses in the phonological match

Choices	Percentage
Pseudowords with matching phonological components	68*
Nonwords with matching phonological components	11
Pseudowords with matching semantic radicals	12
Nonwords with matching semantic radicals	8

* correct response

Children in all age groups showed a preference for "pseudowords with matching phonological components". All the children performed impressively well. However, upon further reflection about this task, it became clear that children might have chosen the "pseudowords with matching phonological components" simply because of the graphic similarity between the choices and the prototypes. In the further analysis, one score was given to the correct choice 'pseudowords with correct phonological component', but no score was given if the same response was chosen again in the semantic match, because that response would be regarded as a match for graphic similarity rather than phonological match. A total of 43% of the responses were affected by graphic similarity.

A one-way ANOVA was conducted to examine the age effect with the number of correct phonological match (revised score) as the dependent variable. No significant age

effect was found, $F(5,30)=1.65$. A series of χ^2 indicated that only the 7- and the 9-year-olds scored significantly above chance level, $\chi^2(3, N=8)=14.55$, $p<0.05$ for 7-year-olds, and $\chi^2(3, N=6)=10.05$, $p<0.05$ for 9-year-olds.

Table 9-2 Mean scores for phonological match by age group

4-year-olds (n=2)	2.0
5-year-olds (n=6)	2.8
6-year-olds (n=6)	3.3
7-year-olds (n=8)	4.4
8-year-olds (n=8)	3.4
9-year-olds (n=6)	4.8

9.1.6.2 The semantic match

Table 9.3 presents the percentage of children's responses to the four choices in the semantic match. Only 26% of the responses made the correct choice - pseudowords with the matching semantic radicals.

Table 9-3 Percentage of children's responses in the semantic match

Choices	Percentage
Pseudowords with matching semantic radicals	26*
Nonwords with matching semantic radicals	13
Pseudowords with matching phonological components	42
Nonwords with matching phonological components	19

* correct match

A further analysis was conducted by giving a point to the correct choice - pseudowords with matching semantic radicals. Table 9.4 presents the mean scores for the choice 'pseudowords with matching semantic radicals' by age group.

Table 9-4 Mean out of ten for choosing correct semantic radicals by age group

4-year-olds (n=2)	0.5
5-year-olds (n=6)	3.3
6-year-olds (n=6)	3.0
7-year-olds (n=8)	1.8
8-year-olds (n=8)	2.3
9-year-olds (n=6)	3.5

A one-way ANOVA was conducted to look at the age effect with correct semantic match as the dependent variable. There was no significant age effect, $F(5,30)=1.60$. A series of χ^2 indicated that none of the age groups performed better than chance in choosing the correct semantic radicals.

9.1.7 Discussion

The study was set up to examine if children could read pseudowords with the help of the phonological and semantic radicals of real words. However, the study failed because of the poor design of the task. The same set of stimuli was used in two sessions. The phonological match was carried out first followed by the semantic match. Although both sections were carried out on different days, older children remembered the set of stimuli and probably their previous choices in the phonological match. When they were asked to perform a semantic match with the same set of stimuli, some of them could not understand why the same stimuli were used again and began looking for different things. For this reason, many of them preferred to stick to the answer they had chosen for the phonological match and picked up “pseudowords with matching phonological components” as the correct responses for the semantic match task.

The setting of the experiment did not arouse the children’s interest. They were confused when one known word (the prototype) and four unknown words (the choices) were presented to them and they were asked to choose an item either similar in meaning or in

pronunciation to the prototype. They could not understand why or how they could match two words similar in meaning or pronunciation when one known word and four unknown words were presented to them. As a result, they could only match for superficial similarity, that being the graphic similarity between the prototype and the choices. The match for graphic similarity was confirmed as 43% of the responses were identical in both the semantic match and the phonological match, which indicated that children could not follow the specific instructions to look for words to match either in meaning or in pronunciation. Furthermore, children were not required to read the pseudowords, instead, they were only required to point to the choices which were likely to have the same pronunciation or the same meaning. In this way, it is difficult to examine if children were adopting any strategies to read novel words.

In the main study, two new tasks would be devised and children's understanding of the phonological and semantic constraints would be examined separately. Children were required to pronounce pseudowords in one task and choose a meaningful pseudoword in another task.

9.2 *The Main Study*

9.2.1 Task 1 - Pronouncing Chinese pseudowords

In this task, children were required to pronounce some pseudowords in a decontextualised situation in order to examine if the phonological information provided by the phonological components is used by children in reading.

9.2.1.1 *Subjects*

One hundred children, aged 5 - 9 participated in this study (for a full description, see Section 5.3.2 in Chapter 5).

9.2.1.2 Design

In this experiment, children were asked to read out the names of the streets on the unknown planet. The names of the streets in the unknown planet were pseudowords which consisted of one character. They were all compound characters with one component being the semantic radical and the other the phonological component. Both components are of frequent use and even five-year-olds should have recognized them without difficulty. The pseudowords were presented in such a way so that children could not rely on either meaning or the situation to predict the pronunciation.

Children were asked to look at the street signs and to try to pronounce the street names one by one. If children were reluctant to pronounce them, four choices (in oral and written form) would be provided for each street name. The choices were made according to the strategies children devised in pronouncing the invented words in Study 3.

1. Strategies based on the phonological component

Children who used this strategy in the previous study tended to read the words by derivation or by analogy. In this study, two choices based on these pronunciation strategies were made:

- by analogy - pronunciation of the word via analogy with other words containing the same non-radical component;
- by derivation - pronunciation of the word derived directly by the pronunciation of the non-radical component;

Choices formed by analogy were made by combining the phonological component of the street name with another semantic radical, and both components were placed in their legal positions. Choices formed by derivation were made by using the phonological component as a single word in isolation. There were two to six possible pronunciations for each pseudoword if children use analogy to pronounce the pseudowords, but since choices were provided, and the pronunciation of the choices

for analogy were entirely different from the pronunciation by derivation, children's preference of the kind of strategies could easily be detected. Children are not expected to get the tones correct because tones are only formed when they are formally used in speech.

2. Strategies based on the semantic component

Choices based on the semantic features were made by combining the semantic radical of the street name with another phonological component, and again both components were placed in their legal positions.

3. The use of nonlinguistic clues to generate the pronunciation

Choices that were used for nonlinguistic clues here did not have the same semantic radical or the phonological component as the pseudowords, but they were also words in frequent use that children were likely to know them well. Some of them were the names of well-known street names in Hong Kong.

Figure 9.2 The street names and its choices

Name of street	by analogy	by derivation	by semantic consideration	by non-linguistic consideration
信	信 /soen ³ /	言 /yin ⁴ /	快 /fai ³ /	他 /ta ¹ /
粒	拉 /lai ¹ /	立 /lap ⁹ /	柱 /tsy ⁵ /	起 /hei ² /
波	波 /bo ¹ /	皮 /pei ⁴ /	妹 /mui ⁶ /	清 /tsing ¹ /
笑	笑 /siu ³ /	天 /tin ¹ /	花 /fa ¹ /	黃 /wong ¹ /
抄	抄 /tsau ¹ /	少 /siu ² /	地 /dei ⁶ /	高 /gou ¹ /

— semantic radical — phonological component — other stroke-patterns

The street names and their choices to be shown to children are always in black, the colour is used here to help the non-Chinese readers

9.2.1.3 Materials

There were five trials altogether. In each trial, the child was presented with a pseudoword and four choices. The pseudowords and the choices were all in handwritten format.

9.2.1.4 Procedure

Children were asked to look at the street signs and to try to pronounce the street names one by one. The street signs were shown in random order. If the children were reluctant to pronounce them, four choices of street name (in oral and written form) were offered.

The children were given the following instructions:

“The planet is a very small one and there are only five streets. But Dung-dung always lost his way because he did not know how to pronounce the street names. Can you help him to pronounce the street names?”

If the children could pronounce the pseudowords, the choices were not shown. If the children had difficulty in pronouncing any one of them, then the choices would be given. The experimenter read out the choices written on a sheet of paper one by one with the following instruction:

“Dung-dung also finds it difficult to pronounce this word, and he cannot decide whether this word should be read as /lai/, /lup/, /shi/ or /hai/? What do you think?”

9.2.1.5 Results

Only 9% of the children produced spontaneous pronunciations and all of them referred to the phonological components for pronunciation. Since the data for spontaneous reading were not sufficient for a separate analysis, they were grouped together with responses in which clues for pronouncing pseudowords were provided. The results is presented in two sections: (1) an overall analysis of the use of phonological components for pronunciation; (2) consistency within child.

9.2.1.5.1 Overall analysis of the use of phonological components

A total of 67% of the responses referred to the phonological components for pronunciation. There were children in all age groups, however, who did not use the phonological clues. They pronounced the pseudowords either by non-linguistic considerations (18%) or by referring to the semantic radicals for pronunciation (16%).

Figure 9.3 presents the percentage of children using different strategies in pronouncing the pseudowords by age group. The 5- and 6-year-olds used all the strategies randomly for pronunciation. At the age of seven, children used more analogy to pronounce words, and began to reject the use of semantic radicals for pronunciation. Eight- and nine-year-olds rejected pronunciation by semantic component and by nonlinguistic consideration, and at the same time, pronounced most of the words by analogy and by derivation.

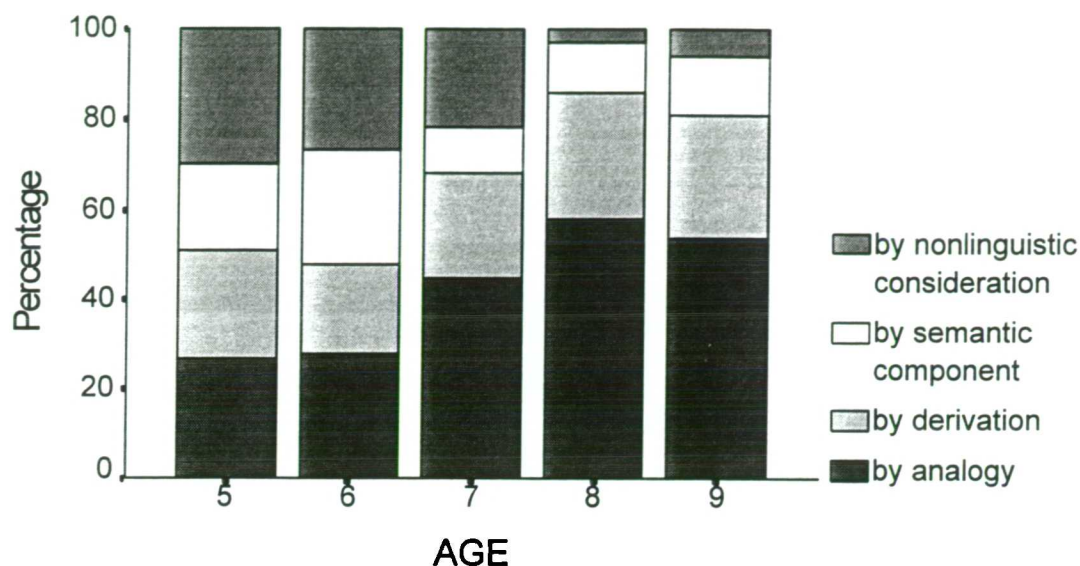
Figure 9.3 Percentage of strategies in pronouncing pseudowords by age group

Table 9.5 presents the frequency of using and not using phonological components for pronunciation by age group. The probability of getting to a correct responses, that is, using the phonological components for pronunciation was 50%. A series of χ^2 indicated that the probability of pronouncing the pseudowords via the phonological components was at chance level for the younger children. Only at age eight do children refer to the phonological components consistently for pronunciation.

Table 9-5 Frequency of using and not using phonological components in pronouncing the pseudowords by age group

Age groups	using phonological components		N	χ^2	df	P
	Yes	No				
5-year-olds	51	49	100	0.80	1	0.37
6-year-olds	48	52	100	0.20	1	0.65
7-year-olds	68	32	100	3.20	1	0.07
8-year-olds	86	14	100	16.20	1	0.00*
9-year-olds	81	19	100	9.80	1	0.00*

* $p < 0.001$

Further analysis was carried out by giving a point to responses based on phonological considerations, which included the use of derivation and analogy in pronunciation. Table 9.6 presents the mean and standard deviation of children using phonological components for pronunciation by age group. A one-way ANOVA was conducted to look at the age effect with the choice of using the phonological components for pronunciation as the dependent variable. A significant age effect was found, $F(4,99)=7.43$, $p<0.001$. Post hoc examination by Scheffe test revealed that the 8- and 9-year-olds scored significantly higher than the 5- and 6-year-olds.

Table 9-6 Mean out of five and standard deviation of children using phonological components for pronunciation

5-years old	2.6 (1.4)
6-years old	2.4 (1.5)
7-years old	3.4 (1.6)
8-years old	4.3 (1.1)
9-years old	4.1 (1.4)
Total	3.3 (1.6)

The two ways of pronunciation via the phonological components were investigated. They were pronunciation by analogy and by derivation. Table 9.7 presents the frequencies of pronouncing pseudowords by derivation and by analogy in different age groups. A series of χ^2 tests indicated that the probability of pronouncing the pseudowords by derivation and by analogy was at chance level for the 5- and 6-year-olds. However, there was a tendency for older children to use analogy rather than derivation in pronouncing pseudowords. The probability of pronouncing pseudowords by analogy was significantly different from chance for the 7-, 8-, and the 9-year-olds.

Table 9-7 Frequency of pronouncing pseudowords by derivation and by analogy by age group

	by analogy	by derivation	N	χ^2	df	P
5-year-olds	27	24	51	0.2	1	0.67
6-year-olds	28	20	48	1.3	1	0.25
7-year-olds	45	23	68	7.1	1	0.00*
8-year-olds	58	28	86	10.5	1	0.00*
9-year-olds	54	27	81	9.0	1	0.00*

* $p < 0.05$

9.2.1.5.2 Consistency within child

Each child was asked to pronounce five pseudowords. Table 9.8 presents the number of children using the same type of strategy in four out of five items. More than 80% of the 8- and 9-year-olds used the phonological components to pronounce the majority of the pseudowords. Not many children used semantic or nonlinguistic considerations in pronouncing the street names consistently, although some younger children indicated a tendency to use nonlinguistic considerations more often than older children.

Table 9-8 Number of children using the same type of strategy for pronunciation in four out of five items

	Phonological consideration			Semantic consideration	Nonlinguistic consideration
	a	b	a or b*		
5-year-olds	2	0	7	0	1
6-year-olds	2	0	5	1	1
7-year-olds	5	0	11	0	2
8-year-olds	6	2	16	0	0
9-year-olds	8	2	17	1	0

a - by analogy b - by derivation

9.2.1.6 Discussion

The study has provided further evidence that children learning to read Chinese can pronounce pseudowords systematically. First, all the children who pronounced the pseudowords spontaneously used the phonological components in the pseudowords to generate the pronunciation. It is a pity that only a few children were willing to do so, and the reason why other children were reluctant to pronounce the pseudowords is believed to be psychological rather than a total lack of linguistic knowledge in this endeavour. Chinese children are not used to the ‘guessing game’ in the process of learning to read.

Second, when children were offered a choice, over 80% of the 8- and 9-year-olds used the phonological components to generate pronunciations for the pseudowords. This was significantly above chance, and they applied the phonological rule consistently in reading four out five items. These results are consistent with the findings of Study 3. Children under the age of eight were at chance level in using phonological components for pronunciation. Younger children prefer nonlinguistic considerations in their pronunciation. They like to pronounce the words in a way similar to the street names they are familiar with, which implies that they have not paid attention to the phonological information provided by the words. Six- and seven years old children prefer semantic considerations for pronunciation. They have paid attention to the kind of information provided by the words but they mistook the semantic radicals as the sound-conveying unit, which is wrong because the phonological information can never be found in the semantic radicals. The performance of the younger children is also consistent with the findings of Study 3.

There are two ways by which pronunciation can be achieved via the phonological components: by derivation and by analogy. Children of all age groups preferred analogy (42%) to derivation (24%). This result is puzzling in view of Chen’s (1993) finding that derivation is the main procedure used by adults for pronouncing Chinese compound words. It is unclear whether the discrepancy is due to a difference in reading

strategies between adults and children, or whether the discrepancy is caused by the confounding effect of graphic similarity between the pseudowords and the choice of pronunciation by analogy in this study. The four choices of pronunciation were shown to children in written form, as it was assumed that it would be too difficult for children to remember the pronunciation by listening alone. Since the words formed by analogy are often more similar in graphic structure to the target word than the phonological component by itself, many children showed a preference for graphically similar words. Further work could aim at finding stronger incentives for children to generate the pronunciation.

The young children in this study could read pseudowords using ‘orthographic-phonology correspondence’ as efficiently as those six- and seven-year-olds in Ho’s (1994) study. In her study, 60% of the six-year-olds in Grade 1 and 70% of the seven-year-olds in Grade 2 could apply the OPC rules to read pseudowords, while in this study, a comparable percentage, 48% of the six-year-olds and 68% of the seven-year-olds could apply the OPC rules. The difference in performance in the six-year-olds may be due to the fact that the children in Ho’s study had read sixty ideophonic compound words before reading the pseudowords, thus the OPC rules were highlighted as the pronunciation of the phonological component matched with the pronunciation of the whole word in all the ideophonic compound words. It is possible that in Ho’s study, some children came to realise the relation between the phonological component and the pronunciation of the word when they were given a chance to read a list of ideophonic compound words where the pronunciation of the phonological component was consistent with the pronunciation of the whole word. Future studies could investigate the effect of practice in acquiring of the OPC rules.

The strong inclination to rely on the semantic radicals for pronunciation in the second task of Study 3 disappeared in this study, as only 16% of the pronunciation referred to the semantic radicals. It confirms the supposition that the semantic radicals were highlighted in the previous task with the presence of the pictures. On the other hand,

there is still a possibility that children prefer to use pronounceable semantic radicals alone to pronounce whole words. Future studies may include the pronounceable semantic radicals as another kind of stimuli.

To conclude, when asked to read some pseudowords in a decontextualised situation, children, especially the 8- and 9-year-olds, were aware of the phonological information embedded in the phonological components. They were able to refer to the phonological components for pronunciation. In the next section, children's understanding of the linguistic functions of semantic radicals is examined.

9.2.2 Task 2 - The semantic constraints in choosing meaningful words

In this experiment, children's conception of how meaning is conveyed in Chinese words was investigated. If children regarded meaning as represented by pictorial depiction, then they would choose an item which consists of illegal graphic patterns but is similar in shape to the object to be represented. If they understood that meaning is conveyed by appropriate semantic radicals, then they would choose a pseudoword which consists of the relevant semantic radicals. Furthermore, the relationship between the semantic function and its position would be investigated again in this experiment.

9.2.2.1 Subjects

One hundred subjects, aged 5 - 9, participated in this study (for a full description of the subjects, refer to section 5.3.2 in Chapter 5).

9.2.2.2 Design

In this experiment, children were to decide the appropriate names for some strange objects on another planet. The strange objects were a crab, a pen, a cup, a horse, and a leaf (Figure 9.4). Children were asked to choose the most appropriate name from the four types of stimuli which consisted of the following characteristics:

1. Pseudowords with correct semantic radicals

The meaning of the pseudowords was represented by the semantic radicals. The semantic radicals provided the appropriate categorization of the objects in the pictures. All the semantic radicals were placed in the semantic radical positions.

2. Pseudowords with incorrect semantic radicals

Although semantic radicals were present, they were irrelevant to and did not convey the appropriate categorization of the objects in the pictures. All the semantic radicals were placed in the semantic radical positions.

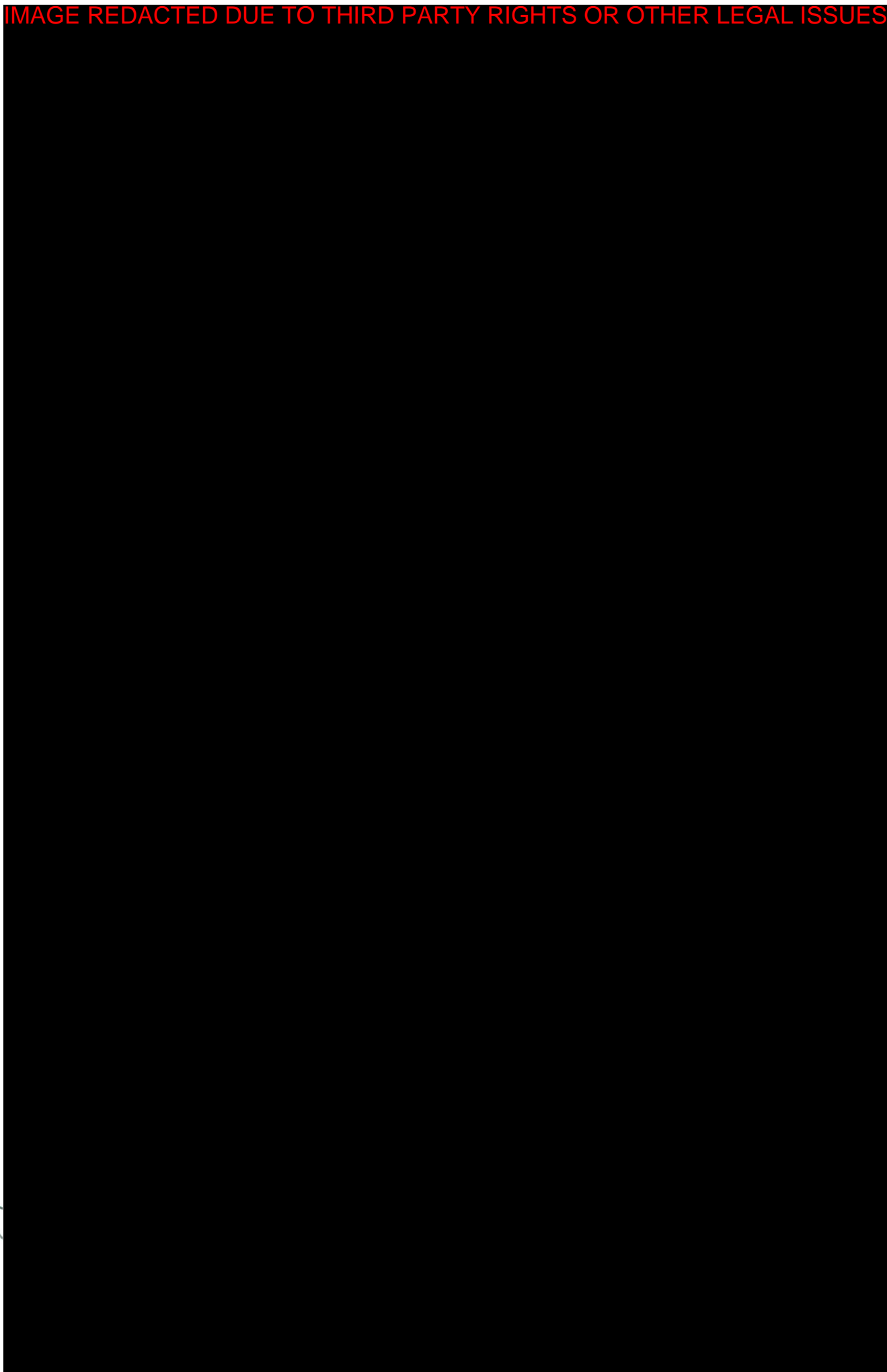
3. Nonwords with illegal positioning of semantic radicals

The semantic radicals contained in the nonwords were appropriate in terms of the categorization of the objects in the pictures. However, they were not placed in the correct position, and therefore could not function as semantic radicals.

4. Nonwords with illegal graphic patterns

They were formed by combining a semantic radical and a pictorial graphic pattern together. Pictorial graphic patterns, which are not used as stroke-patterns, were used in these nonwords. They were designed in order to bear some graphic similarities with the objects in the pictures, e.g. the nonword for *pen* < > used the shape of the pencil, and the nonword for a strange *animal* < > depicts the eyes and legs of the donkey.

IMAGE REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



9.2.2.3 *Materials*

There were five trials altogether. In each trial, children were presented with a colourful picture of a strange object and four choices printed on a A4 size paper. The choices were scattered in different locations on different pages.

9.2.2.4 *Procedure*

In this experiment, children were told that Dung-dung was going to decide the appropriate names for some strange objects in another planet. Children were given the following instruction:

“In another planet, Dung-dung began to teach the spacemen to write Chinese words, and some children have made up words for the strange objects that they have found on the planet. Which one do you think is the most appropriate name for the objects?”

9.2.2.5 *Results*

Children were asked to choose from four choices a pseudoword which was the most appropriate name for some strange objects in the pictures. Table 9.9 presents the frequencies of children's responses to the five objects. A series of χ^2 indicated that there was no significant difference in the children's choice of stimuli in different objects, except in choosing 'nonwords with illegal positioning'. Children chose significantly more 'nonwords with illegal positioning' in the pictures *horse* and *crab*, probably because these semantic radicals do not have fixed graphic positions.

Table 9-9 Frequency of children's responses to the five pictures

Choices	Objects					N	χ^2	df	P
	horse	cup	leaf	crab	pen				
1.	14	20	28	18	20	100	2.60	4	0.63
2.	37	17	15	35	12	116	14.02	4	0.01*
3.	8	28	12	10	27	85	9.40	4	0.05
4.	41	35	45	37	41	199	1.52	4	0.82

1. nonwords with illegal graphics

2. nonwords with illegal positioning

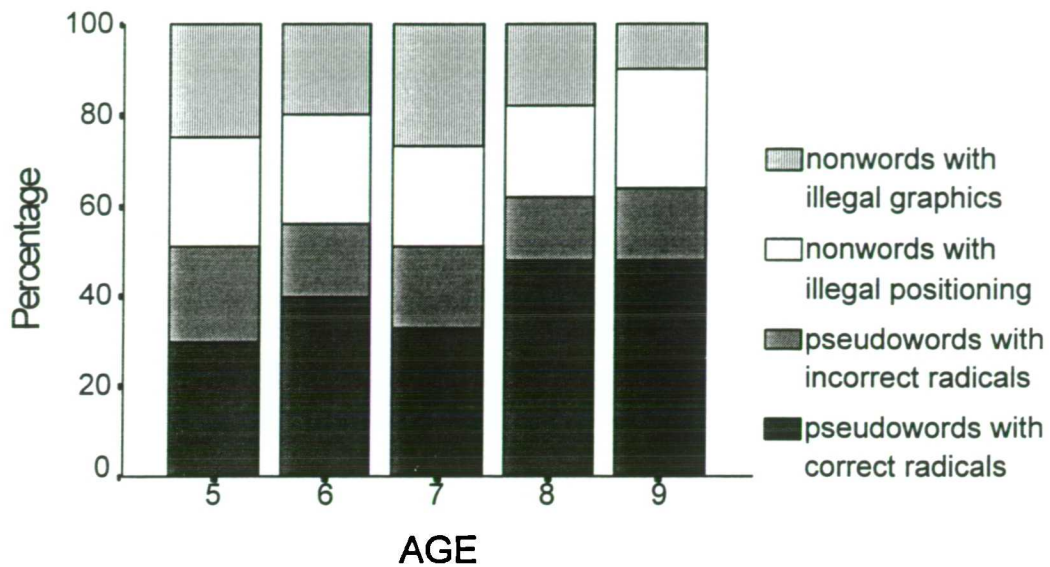
3. pseudowords with incorrect radicals

4. pseudowords with correct radicals

* $p < 0.05$

Figure 9.5 presents the percentage of choice for different types of stimuli by age group. The highest proportion, 40%, of the responses fell on 'pseudowords with correct semantic radicals'. With four choices, the chance level of choosing the correct choice 'pseudowords with correct semantic radicals' was 0.25. The frequency of selection of the different type of stimuli was significantly different from chance for the 6-, 8- and 9-year-olds, $\chi^2(3, N = 20) = 8.96$, $p < 0.05$ for 6-year-olds, $\chi^2(3, N = 20) = 13.90$, $p < 0.05$ for 8-year-olds, $\chi^2(3, N = 20) = 17.40$, $p < 0.05$ for 9-year-olds.

Furthermore, a series of χ^2 was also conducted to examine the probability of choosing 'nonwords with illegal graphic patterns' in different age groups. No child from any age groups chose 'nonwords with illegal graphic patterns' significantly more often than expected by chance.

Figure 9.5 Percentage of choice of word types by age group

Further analysis was carried out by giving a point for each choice of correct semantic radical. A one-way ANOVA was conducted to look at the age effect with the choice for 'appropriate pseudowords with correct semantic radicals' as the dependent variable. A significant age effect was found, $F(4,99)=2.76$, $p<0.05$. Post hoc examination by LSD test showed that the eight- and nine-year-olds made more correct choices than the five- and seven-year-olds, but the significance was marginal because no significant difference was found between the mean scores of various age groups by the more stringent Student-Newman-Keuls test. A developmental trend was traced with older children being able to consider both orthographic and semantic information in deciding appropriate names for different objects, while younger children overlooked one kind of information or the other.

Each child was asked to choose five names for the objects. Table 9.10 presents the number of children making the same choice of word types in four out of five items. Not many children chose the same word type consistently. A varied, rather than

uniform pattern of choice was found in almost all age groups. 8- and 9-year-olds were able to make more consistent choice of the 'pseudowords with correct radicals'.

Table 9.10 *Number of children choosing the same word type in reading in four out of five items*

	nonwords with illegal graphics	nonwords with illegal radicals	pseudowords with incorrect radicals	pseudowords with correct radicals
5-year-olds	0	0	1	0
6-year-olds	0	1	1	1
7-year-olds	1	2	0	0
8-year-olds	0	0	0	3
9-year-olds	0	0	2	5

9.2.2.6 Discussion

When the children were asked to look for a pseudoword as a suitable name for a particular object, they had to find the connections between the object and the choice of stimuli. There were two kinds of connections, a linguistic connection and a pictorial connection. Children have already acquired the knowledge that Chinese words are represented by semantic radicals. The 6-, 8- and 9-year-olds were able to choose the 'pseudowords with correct radicals' significantly more often than chance, and a significant age trend was traced which indicated that older children were more advanced in understanding the linguistic function of the semantic radicals.

The study also provides evidence that even 5-year-olds do not regard the meaning of the Chinese words as presented by pictographs. None of the age groups selected the 'nonwords with illegal graphic patterns' significantly above chance level. The result of this study was comparable to the findings in Study 2 (Orthographic acceptability of pseudowords and nonwords) where 'nonwords with illegal graphic features' were rejected by all age groups. Thus it is concluded that even with the effect of graphic

similarity between the objects and the words, children could still reject nonwords with illegal graphic patterns.

This finding however contradicts the result in Chan's (1990) study where preschool children drew pictures in their writings. The difference may be caused by age, as children in this study are older; and it may also be a difference in the task demand. Although children know that pictures are not words, they believe that they can also be used to convey meaning.

To sum up, children at the age of six begin to be aware of various semantic constraints including the type of semantic representation, the positioning of semantic radicals, and the meaning conveyed by different semantic radicals. To combine all this knowledge in order to decide the appropriate word to convey meaning is not a simple task. And although the six-year-olds have indicated a preference for 'pseudowords with correct radicals', they cannot integrate all these constraints and choose the correct response across all items. Only the eight- and nine-year-olds can choose the correct responses 'pseudowords with correct semantic radicals' consistently.

10. Chapter 10

Conclusion and Discussion

The thesis begins with a presentation of two entirely different viewpoints with regard to the word recognition process in Chinese. The ‘holistic’ approach’ treats characters as logographs. A character is believed to be the smallest representational unit, and in order to be able to read, a Chinese reader has to retrieve lexical information directly from memory, which contains at least two to three thousand characters (Cheng, 1992; Perfetti & Zhang, 1991, 1995). The ‘analytical approach’ regards each Chinese character as a combination of stroke-patterns, and the stroke-patterns provide linguistic information which facilitates learning. These two different approaches, one based on rote memory and the other based on linguistic analysis are incompatible. If one is proved to be correct, the other approach is destined to be wrong.

Recent psycholinguistic research has highlighted the importance of the stroke-pattern in reading Chinese (Chen, 1993, 1995; Flores d’Arcais, 1992, 1995). Instead of viewing individual characters as distinct configurations with direct access to meaning and pronunciation, the stroke-patterns provide orthographic, semantic and phonological information at an intra-character level. Evidence has been gathered to support the view that adult readers rely on the semantic and phonological information conveyed by the stroke-patterns in the process of recognizing a Chinese word (Zhang & Simon, 1985; Seidenberg, 1985; Zhu, 1987; Hue & Erickson, 1988; Flores d’ Arcais, 1992, 1995; Hue, 1992; Chen, 1993, 1995). The ability to utilize the orthographic, semantic and phonological information provided by the stroke-patterns has only been examined in fluent Chinese adult readers. No study has yet investigated if written Chinese is learned analytically from the beginning of the process of learning to read.

The present study adopted an analytical approach to investigate the nature of learning to read and write Chinese. In Section 1, an overview of the findings of the experiments is presented. The results of the present study suggests an entirely different view with regard to the nature of learning to read and write Chinese. In line with the findings of recent psycholinguistic research, it is demonstrated in this study that children learn to read and write Chinese by acquiring the proficiency in analyzing the orthographic and linguistic information at an intra-character level. They use stroke-pattern as the unit of visual analysis which facilitates their speed and accuracy in reading Chinese words. Children judge the orthographic acceptability of Chinese words according to the graphic characteristics and positioning of the stroke-patterns, and rely on the linguistic components - the semantic radicals and the phonological components - for meaning and pronunciation. Based on these findings, it is believed that written Chinese is learnt analytically rather than by rote memory.

In Section 2, an outline of a developmental theory of learning to read and write Chinese is proposed, in an attempt to encourage further research in this direction. It is suggested that children progress from global to specific - first by developing global graphic and phonological awareness, and then proceed to analysis the intra-character units. The intra-character analysis, that is understanding the formal and functional constraints of the stroke-patterns, may be an important predictor in children's ability to learn to read. Children are likely to progress through developmental stages, each stage is marked by children being able to acquire a deeper understanding of written Chinese.

In Section 3, some suggestions for teaching implications in light of the new findings from this study are presented. The present teaching methodology in teaching written Chinese is not designed to facilitate children to acquire an understanding of the language at an intra-character level. Teachers are encouraged to examine children's learning process of written Chinese from a different perspective.

10.1 Process of learning to read and write Chinese - holistic versus analytical

In this section, an overview of the findings from a series of experiments is presented. These experiments were set up to investigate if linguistic information provided at an intra-character level is utilized by children in the beginning stage of learning to read and write. Children were asked to read and write some pseudowords and their ability to analysis the orthographic, semantic and phonological information at an intra-character level was examined. If it is true that written Chinese is learnt simply by rote, it is impossible for any Chinese children to read pseudowords, as no lexical information is provided by this type of graphic patterns; and it is impossible to trace any systematic patterns in their creative spelling. If learning to read Chinese requires a deeper level of understanding, and not by memorizing the characters, then at least three kinds of evidence can be found at an intra-character level. First, children use the stroke-patterns as the basic perceptual unit in word recognition. Second, the stroke-patterns pose formal and functional constraints in early reading. Third, the knowledge of the formal and functional constraints of the stroke-patterns is acquired gradually, and not in an all or nothing fashion.

A variety of tasks were set up which included judgment tasks, spelling tasks, decoding, and reading for meaning tasks. If children have acquired an understanding of written Chinese at a cognitive level, they should be able to demonstrate their knowledge in different kinds of task demand. The formal characteristics of written Chinese: its positioning and graphic features, were investigated by the judgment tasks and the writing tasks. The functional characteristics of its linguistic units - the semantic radicals and the phonological components were investigated in both the reading and writing tasks. The cues involved in reading and writing can be quite independent (Bryant and Bradley, 1980; Frith, 1980) and there was a possibility that children could only attend to the linguistic functions in either one of the tasks.

10.1.1 Unit of visual analysis of Chinese words: the stroke-patterns

A visual discrimination task was set up to examine if children use strokes or stroke-patterns as the unit of visual analysis in Chinese words. Children were presented with pairs of pseudowords which differed in the number of strokes and the number of stroke-patterns, and they were asked to decide as quickly as possible whether a series of pseudoword pairs were the same or different. It was found that children have less difficulty in differentiating words pairs with fewer stroke-patterns than word pairs with more stroke-patterns, with the number of strokes remained constant. Five-year-olds are fairly accurate in differentiating word pairs with fewer stroke-patterns, but they cannot perform equally well in judging word pairs with a larger number of stroke-patterns. Nine-year-olds are more competent in discriminating word pairs than younger children, although they, too, are more accurate in judging words pairs with fewer stroke-patterns than word pairs with more stroke-patterns.

Thus, like the adult readers in Chen's (1995) study, children at the age of five already rely on stroke-patterns as the basic unit in word recognition. The presence of visual units at an intra-character level is confirmed. In the next section, I will present evidence which showed that these stroke-patterns do not function as visual units merely, they also function as orthographic and linguistic units, and children can utilize these functions in reading and writing.

10.1.2 The formal constraints of written Chinese - the positioning and graphic characteristics of the stroke-patterns

The formal constraints of written Chinese - the positioning and graphic characteristics of the stroke-patterns - were addressed in two studies. The first study was an orthographic acceptability task where children were required to differentiate pseudowords from nonwords, on the basis of the positioning and graphic

characteristics of the stroke-patterns. The second one was a writing task in which children were asked to create some pseudowords. Clear evidence was found that children follow the positional and graphic rules in both the judgement and production tasks.

In the orthographic acceptability task, it was found that children judge the orthographic acceptability of pseudowords and nonwords on the basis of the positioning of stroke-patterns and their graphic characteristics. Children at the age of five begin to discriminate pseudowords from nonwords, and the ability to accept pseudowords and reject nonwords increased with age. While 5-year-olds demonstrate only a slight preference for pseudowords over nonwords, 9-year-olds make a clear distinction in their preference for pseudowords over nonwords.

Children also demonstrated their knowledge of the legal positioning of stroke-patterns in the creative spelling tasks. They were asked to form pseudowords by combining the stroke-patterns together in any way they liked. Instead of placing the stroke-patterns at random and making nonwords which violate the orthographic rules of written Chinese, most of the children placed the stroke-patterns in their legal positions and formed pseudowords. Five-year-olds already demonstrated an understanding of the orthographic rules by forming some pseudowords. The ability to form pseudowords according to the orthographic rules increases with age. Six-year-olds formed more pseudowords and nine-year-olds spelled almost all the pseudowords correctly by placing the stroke-patterns in their appropriate positions. The results from the two experiments allow us to conclude that children at the age of five have acquired the knowledge of the legal positions and the graphic characteristics of the stroke-patterns in the orthographic structure of written Chinese. This knowledge helps children discriminate pseudowords from nonwords, and helps them create pseudowords in their writing.

10.1.3 The functional characteristics of the stroke-patterns

10.1.3.1 *The semantic radicals*

In the creative writing tasks, children were asked to make up some names for strange objects. They were given different kinds of semantic radicals in two tasks. In the first task, children were provided with semantic radicals with fixed graphic positions, and in the second task, they were given semantic radicals which could be placed in some other positions if they were not used as semantic radicals. Children had to follow the orthographic rule to place the semantic radicals in the appropriate position in the first task, and applied the semantic radical position rule to place the semantic radicals in the second task.

Most of the children understood the function of semantic radicals. They matched the semantic radicals to the appropriate objects, in order to represent the category of the objects in the pseudowords. It appears that 5-year-olds do not yet understand the function of semantic radicals; 7-year-olds attempt to match the familiar semantic radicals to the appropriate objects; and 9-year-olds can use the semantic radicals to represent meaning quite efficiently. In addition, it is demonstrated in Task 2 that 8- and 9-year-olds can decide the positioning of semantic radicals purely by their function. When they are presented with stroke-patterns which have no fixed graphic positions, they decide the linguistic functions of the stroke-patterns, and for stroke-pattern in which the linguistic function is semantic, they rely on the semantic radical position rule to place the semantic radicals.

In the reading task, the children were asked to choose a pseudoword from among four items to be the most appropriate name for a strange object. Some of the items used pictographs to represent meaning and some of them used semantic radicals. Children regard the semantic radical as the only representational unit for meaning in

Chinese words. Even 5-year-olds do not regard meaning as conveyed in pictographic form in written Chinese. None of the age groups chose the word type 'nonwords with illegal graphics', in which the special features of the objects were represented by drawing, beyond chance level. Eight and nine-year-olds relied on both the semantic information provided by the semantic radicals and their positioning to choose the most meaningful words. The frequency of their choice of 'pseudowords with the correct semantic radicals' was significantly different from chance.

10.1.3.2 The phonological components

Two types of evidence were gathered in examining children's understanding of the functions of the phonological components. Children were asked to pronounce their own spelling, and to read pseudowords in a situation where they could not refer to specific meaning (that is, the pseudowords were presented as street names on an unknown planet in the task).

In the creative spelling tasks, children were asked to invent written names for peculiar objects and read them in two tasks with different levels of complexity. In the simpler task, all the semantic radicals were unpronounceable and the phonological components were the only pronounceable units. In the more complex task, both the semantic radicals and the phonological components could be pronounced. In the simpler task, a total of 33% of the pseudowords were read by referring to the phonological components. An upward trend was observed with older children pronouncing more words via the phonological component than younger children. It appears that 5-year-olds do not use a reading strategy which is based on the phonological component whereas 9-year-olds seem to rely on the phonological information provided by the phonological component for the majority of their attempts to read pseudowords.

In the more complex task, two pronounceable components were provided. Children had to decide between the semantic radical and the phonological component in a pseudoword which was the appropriate phonological unit. Most of the children consistently referred either to the semantic radicals or the phonological components for pronunciation. The younger children preferred to pronounce the semantic radicals. Only the 9-year-olds relied on the phonological components for pronunciation. Therefore, it is suggested that Chinese children are aware of the linguistic function of the phonological components. However, this knowledge appears in a later, rather than an early stage of learning to read. In the present study, only the 9-year-olds could refer to the phonological components systematically in pronouncing the pseudowords.

In the reading task, the situation was set up in a way similar to reading novel words out of context, as children were asked to read five pseudowords which were the names of the streets on an unknown planet. Children responded to the pseudowords in the same way as they pronounced their own creative spelling. The younger children disregarded the phonological components while the older children consistently referred to the phonological components for pronunciation. Five-, six- and seven-year-olds used different strategies to pronounce pseudowords, and the 8- and 9-year-olds referred solely to the phonological components for pronunciation.

There are two ways of pronunciation via the phonological component: analogy and derivation. There is a tendency for children to use analogy rather than derivation in pronouncing pseudowords. Seven-, eight- and nine-year-olds pronounce the pseudowords by analogy predominantly. The findings in the reading tasks lent further support to the result of the creative spelling tasks that 8- and 9-year-olds pronounce Chinese novel words by systematically referring to the phonological components for pronunciation.

To conclude, evidence has been gathered in all the experimental studies that children analyze Chinese script at an intra-character level. It is emphasized that all the tasks required the children to use their implicit knowledge of the system of written Chinese, because they were required to read and write pseudowords, which have no lexical representation; and the children could not rely on their knowledge of real words to perform any of the tasks. If children had not developed any understanding of the underlying rules of written Chinese, it would be impossible to trace any kind of development in either the form or the functions of the linguistic components. The results from this study offer an entirely different picture of how children learn to read Chinese, and allow one to conclude that Chinese children do not learn the Chinese script simply by rote, but require an understanding of the linguistic information provided at an intra-character level.

10.2 The outline of a developmental theory of children learning to read and write Chinese

Since we no longer believe that children learn to read and write Chinese by merely memorizing Chinese characters one by one, then what are the essential elements that children have to learn in order to be literate? What is the process of children learning to read and write Chinese? In the following section, I would like to propose an outline of a developmental theory of children learning to read and write in Chinese. The proposal is basically hypothetical, although the suggestions are based on the results of my study.

The suggested developmental theory is not based on global graphic and phonological awareness, and the insufficiencies of these skills to account for the entire learning process are presented. It is suggested that one of the essential elements that children have to learn to read Chinese is to develop an ability to analysis intra-character orthographic and linguistic information. Intra-character analysis may be the best predictor in developing reading ability in Chinese, as findings from previous research has pointed to this direction. The understanding of the orthographic and linguistic information at an intra-character is possibly developed in stages, each stage being marked by a deeper understanding of the underlying rules of written Chinese.

10.2.1 The global identification of the graphic and phonological units

Chinese children begin their search for the nature of the representational system of written Chinese at an early age. Preschool children apply unique graphic features exclusive for written Chinese in their creative spelling. Three-year-olds use different graphic representation for drawing and writing, and at the age of five, they create their own writing with stroke-patterns which are constricted in size and square in shape (Chan, 1992). Preschool children are aware that words convey meaning. They demonstrate high semantic intent when they read environmental print and they gradually establish a representational relationship between print and the referent item (Lee, 1989). Children at preschool level have also acquired the knowledge of one-to-one correspondence between the syllable and the character in reading and writing. Five-year-olds apply syllabic rule in reading unknown words (Chan, 1990), and use the correct number of characters in writing their names (Chan, 1992).

At preschool stage, Chinese children have acquired the basic understanding between script and speech, and can differentiate written Chinese from other graphic patterns. However they have not yet acquired sufficient knowledge to be independent readers. For instance, they need the support of context to help them to guess the meaning of the written words, and if the words are presented out of context, they have no cue to enable them to read (Lee, 1989). In other words, they have not yet acquired the knowledge to analyse the words to look for cues that can convey meaning. Although they have acquired the form of written Chinese at an early stage, most of the words they wrote were approximate writing (Chan, 1992). The words that they have invented do not follow the orthographic and linguistic structure of written Chinese, and they can neither provide meaning nor pronunciation. The acquisition of the one-to-one correspondence between the syllable and the character is clearly not sufficient for reading.

10.2.2 Intra-character analysis

Over 90% of the Chinese characters are compounds and they are formed by combining stroke-patterns together. These stroke-patterns provide visual, orthographic and

linguistic information. Intra-character analysis refers to paying attention to the stroke-patterns, and not only to the character as a whole. There are at least two reasons which account for the need to develop intra-character analysis in the process of learning to read Chinese. First, although children can learn some characters purely by visual memory, they begin to face difficulties in memorizing more and more characters. They have to devise a mechanism to help them to approach the characters in a systematic way. Second, when faced with a novel word, they have to develop some strategies to search for its meaning and pronunciation.

How can intra-character analysis facilitate memory?

It is demonstrated in this study that even 5-year-olds are very good in Chinese word discrimination. They have not difficulty differentiating pairs of novel words with many strokes. The reason which accounts for this is probably because children are familiar with the stroke-patterns of the novel words. The stroke-pattern is taken as a whole in visual comparison, and instead of comparing plenty of strokes, they are comparing 2 - 3 stroke-patterns, and it accounts for the reason why children can compare a pair of words within a few seconds. Using the stroke-pattern as the basic visual unit therefore, facilitates the word recognition process.

The stroke-pattern also functions as an orthographic unit. Children decide whether an unknown word can be acceptable as a Chinese word according to two criteria: the positioning and the graphic characteristics of the stroke-patterns. Children at the age of five begin to reject nonwords which consisted of illegal graphic patterns such as circles and illegal shapes. They develop some assumptions that some graphic patterns cannot be acceptable in written Chinese. They can also make orthographic decisions on the basis of the legal positioning of stroke-patterns. Older children improve significantly in deciding the orthographic acceptability of graphic units, and this skill is believed to be an important skill to be developed in order to master the orthographic structure of written Chinese.

The knowledge of the legal positioning of stroke-patterns not only enables children to identify 'wrongly spelt' words, but it also guides them to spell words correctly. Six-year-olds can allocate the stroke-patterns to their respective locations and form pseudowords, instead of nonwords. The ability to relate a stroke-pattern to its position is likely to facilitate the learning of new character, as the retrieval of the stroke-pattern is automatically accompanied by the allocation of the stroke-pattern to the possible position.

How can intra-character analysis help in decoding new words?

When faced with a new character, children rely on the linguistic rules of written Chinese to understand and pronounce it. Seven-year-olds can refer to familiar semantic radicals with fixed positions for meaning; and eight-year-olds can refer to semantic radicals with or without fixed positions for meaning. Looking for semantic radicals is assisted by the graphic rule and the positional rule. Some semantic radicals are easy to identify because they have fixed positions; and some are more difficult because they can also function as phonological components when they are placed in different positions. The knowledge that most of the semantic radicals can be found on the left or on top of characters can be helpful in identifying

some of the semantic radicals. However, confusion still arises as some semantic radicals can also be found on the right or at the bottom, too, and it took some time for children to work out the exceptions.

It is not difficult to identify the phonological components, as they can usually be found on the left or in the bottom of the words, and they are usually the more dominant graphic unit in the characters. There are two ways in which pronunciation can be derived via the phonological components: by derivation, if the phonological component is pronounceable, and/or by analogy, if the phonological component is non-pronounceable. It was found that only 9-year-olds referred consistently to the phonological components for pronunciation. Since there are many Chinese characters where the phonological components are not used for pronunciation, children fail to discover the relationship between speech and print in an intra-character in their early stage of learning to read.

The greatest confusion that Chinese children have to face when they begin to analyse the intra-character information probably lies in the differentiation between ideophonic compound words and ideocompound words. The functional characteristics of the stroke-patterns in these two types of words are very different. In ideocompound words, the meaning of a character is achieved by combining the different meanings conveyed by the stroke-patterns, and there is no phonological component in any ideocompound words. In ideophonic compound words, one component conveys meaning and the other one conveys sound.

When faced with a novel word, it is difficult to decide which type of word it is. In this study, we have evidence that children produce compound words and regard them as simple characters, and they failed to use the stroke-patterns to convey either the meaning or the pronunciation. Some children made ideocompound words, but

were not sure how to pronounce them, and they used the semantic radicals for pronunciation. Findings from adult research have provided evidence that when faced with a new word, fluent Chinese readers refer to the semantic radicals for meaning and phonological components for pronunciation. In other words, when they are faced with a new word, they tend to treat it as a ideophonic compound word, and then systematically resolve to reach for meaning and pronunciation. Findings from this study also indicated that older children are using this strategy to read novel words. Based on this finding, it is suggested that treating the new words initially as ideophonic compound words may be a more efficient way of learning to read; first because there are more ideophonic compound words than the other two types of words, second, it is the only consistent strategy which children can address to meaning and sound in reading Chinese.

10.2.3 Predicting the development of reading ability in Chinese

It is possible that intra-character analysis is the best predictor of children's ability to read Chinese. Children rely on stroke-patterns for visual discrimination in the beginning stage of their learning to read. Then they refer to the orthographic and linguistic information provided by the stroke-patterns to read and write Chinese. Stroke-patterns play an important role in word recognition and enable children to decide quickly whether some words are wrongly spelt. At the same time, the knowledge of the presence of the stroke-patterns also help children to learn to write systematically as they come to know the stroke-pattern in relationship to its location. In other words, when children are presented with stroke-patterns, they know roughly where they should be placed in order to form a Chinese character. Children can also utilize the linguistic functions of the stroke-patterns to decode new words, by referring to the semantic radicals for meaning and the phonological components for pronunciation.

Furthermore, other evidence has also suggested that the awareness of the stroke-patterns is likely to have a facilitative effect in the process of learning to read Chinese. Ho (1994) investigated the kind of global phonological and visual awareness which predict children's ability to read, and she found the Frostig Developmental Test of Visual Perception especially Subtest III (Constancy of Shape) to be the most significant and consistent predictor of children reading Chinese. She believes that the importance of the shape constancy ability for reading Chinese is probably due to the fact that children have to detect recurring stroke-patterns, which are basic constituents of Chinese characters.

On the other hand, phonological awareness may not be a significant predictor in order to learn written Chinese. In the present study, only the 8-year-olds could refer to the phonological components systematically for pronunciation, and by that time, they had already learnt about 2,000 characters.

Since over 90% of the characters are formed by combining stroke-patterns together, the awareness of the stroke-patterns is likely to have a facilitative effect in the process of character recognition. However, there is a lack of evidence in how this knowledge facilitates text reading. Chen, Lau and Yung (1993) asked first to sixth graders in Hong Kong to search for a stroke-pattern *mouth* < □ > in both meaningful and not meaningful texts, and found that children generally perform very well in detecting the stroke-pattern. However, the first and the second grade children actually did better than the third, fourth and fifth grade children. This suggests that paying attention to stroke-patterns may be particularly crucial for the first and second graders when their main priority is to learn characters. Later on, when they have learned a sufficient amount of characters, they may shift their strategies, and no longer rely on stroke-patterns in text reading.

The belief that intra-character analysis as the best predictor of children's ability to read is purely hypothetical and subject to inspection. Future research into the nature of the learning process involved in reading and writing Chinese can examine the following assumptions. First, if intra-character analysis is not the best predictor of children's ability to read Chinese, then what is a stronger predictor which facilitates the written Chinese learning process? Second, if intra-character analysis is the best predictor of children's ability to read Chinese, is it also a criterion in order to differentiate good readers from poor readers? What is the relationship between the ability to analyse intra-character information and the ability to read Chinese? In addition, how does the ability to analyse intra-character information develop? A model of the possible developmental stages of intra-character analysis is proposed in the following section.

10.2.4 Proposed developmental stages in understanding written Chinese at an intra-character level

Children begin to pay attention to stroke-patterns at the age of five, and gradually develop various degree of understanding of their orthographic and linguistic functions. The development is marked with some signposts which characterize the cognitive conflicts which are faced by children in the process of learning to read Chinese. Five developmental levels, based on the findings from this research, are proposed. Each level is illustrated with examples of individual children responding to the orthographic acceptability task, the creative writing tasks, and the pseudoword reading tasks.

Level 1

Being able to read a few characters does not mark the beginning of an understanding of the form and functions of the stroke-patterns. Level 1 is characterized by an inability to utilize the orthographic and linguistic information provided at an intra-

character level to facilitate read and writing. Li (5 years 1 month) has only recently been introduced to written Chinese. He knew some of the Chinese characters like *car* < 車 >, *fish* < 魚 >, and *woman* < 女 >. However, in the orthographic acceptability task, he could not follow any orthographic rules to differentiate pseudowords from nonwords, and he judged equal number of pseudowords and nonwords as acceptable ‘words’. When he was asked to create some writing, he could not form any pseudowords, he allocated the stroke-patterns at random with no consideration of the correct positioning of the stroke-patterns.

He failed to acknowledge that semantic radicals convey meaning. In the creative spelling tasks, he could not match most of the semantic radicals to the corresponding pictures. When he was asked to read some pseudowords for meaning, he was not able to choose the pseudowords which consisted of the correct semantic radicals and conveyed the desired meaning.

There is a one-to-one correspondence between a character and a syllable in Chinese, and Li had not yet acquired this knowledge. In pronouncing the pseudowords that he had created, Li did not match the pronunciation on the syllable level, he pronounced a pseudoword < 女虫 > which consisted of only one character with three syllables [tai-toi-dum]. Furthermore, he did not refer to the phonological components to pronounce the pseudowords in the reading tasks.

Level 2

Level 2 is marked by an initial awareness of the presence of stroke-patterns. A lot of confusion arises in understanding both the form and the functions of the stroke-patterns. Lo (6 years 0 month) had been introduced to written Chinese informally in kindergarten for more than one year. He could recognize about a hundred characters and could write about 30 - 40 of them. However, when he was asked to differentiate

some pseudowords and nonwords in the orthographic acceptability task, he did not differentiate them very well. He accepted quite a lot of pseudowords and nonwords with illegal graphic features, although he also indicated a slight rejection of nonwords with illegal positions. His inclination to accept nonwords with illegal graphic features were more obvious as he chose 60% of the nonwords with illegal graphic features as names of some strange objects.

Lo was aware of the semantic function of the stroke-patterns. In the reading task where he was asked to choose the most suitable name to represent an object, all the words he chose had the correct semantic radicals in them. The knowledge of presence of semantic radicals also helped him to use semantic radicals correctly in the creative spelling tasks.

Although he had some knowledge of the presence of semantic radicals, he was confused about their function and regarded the semantic radicals as the source of pronunciation. He pronounced 80% of the pseudowords by semantic radicals in both the reading task and the creative writing tasks.

Level 3

Level 3 was marked by a gradual improvement in handling the formal constraints of the stroke-patterns, such as their graphic structure and positional rules. Ma (7 years 0 month) attended to the positional and graphic characteristics of the stroke-patterns. She rejected 90% of the nonwords with illegal graphic features and 50% of the nonwords with illegal positions, but accepted 80% of the pseudowords in the orthographic acceptability task. In the reading tasks, she again rejected all the nonwords with illegal graphic features and nonwords with illegal positioning of stroke-patterns.

Although Ma was able to form pseudowords most of the time in the creative spelling tasks, she was confused about the positioning of the semantic radicals when the stroke-patterns have no fixed positions. She pondered for quite a while and shifted the position of the stroke-patterns several times. She had not yet acquired the knowledge that most of the semantic radicals are found either on top or on the left of words, and she made two mistakes by placing the semantic radicals in the non-radical positions. Furthermore, she did not pay enough attention to the meaning of the semantic radicals, and chose two pseudowords with inappropriate semantic radicals in the reading tasks.

In pronouncing the pseudowords, Ma did not use any of the nonlinguistic consideration for pronunciation. She knew that words provided cues for pronunciation but she was not sure where the cues were. Sometimes she used the semantic radicals for pronunciation and sometimes she used the phonological components for pronunciation. She had not yet acquired the knowledge that semantic radicals can never be pronounced in any compound words because in Chinese, semantic and phonological information are encoded in separate orthographic components.

Level 4

This level is marked by the awareness of the relationship between the positions of the stroke-patterns and its functions. Tang (8 year 4 months) had no difficulty in differentiating pseudowords from nonwords in the orthographic acceptability task. She could place all the semantic radicals in the appropriate place in her creative spelling. When she made the pseudoword as the name for a strange fish, she said, “I am going to put this stroke-pattern *fish* < 魚 > on this side (left hand side). Although this *fish* < 魚 > can also be put on this side (right hand side), it is not right for this word.” Tang noticed that the stroke-pattern *fish* < 魚 > can be found on

both the left and the right side of words. However, when it is used as a semantic radical, it is always on the left hand side.

In pronouncing the pseudowords, Tang rejected pronunciation by nonlinguistic consideration. Like Ma, she knew that pronunciation cues could be found in the words but she was not sure where they were. Thus, she used both the phonological components and the semantic radicals for pronunciation. Again she had not yet acquired the knowledge that semantic radicals are never pronounced in Chinese.

Level 5

This level is marked by an understanding of the formal and functional constraints of the stroke-patterns, and being able to refer to the phonological components for pronunciation. Chan (9 year 1 month) rejected all the nonwords with illegal positioning of stroke-patterns and 90% of the nonwords with illegal graphic features. At the same time, she accepted 80% of the pseudowords as words, implying that she made her judgement based on orthographic acceptability, rather than by lexicality. She used the semantic radicals to represent meaning, and was able to place all the semantic radicals in their legal positions. Furthermore, she used the phonological components for pronunciation consistently in pronouncing all the words.

The proposed developmental stages of intra-character analysis, however, is not intended to be used as a stage model, it can only be regarded as signposts which marked a more advanced level of understanding of written Chinese. Nunes, Bryant, and Bindman (Submitted for publication) have clearly stated that any stage model should pass at least three tests. One is that all the children in the appropriate age range should clearly belong to one of the stages in the model. A second requirement is that children at more advanced stage of the development should be older or

educationally more successful. The third is that individual children over time should move from less to more advanced stage, but not in another direction. A longitudinal study to verify the proposed developmental framework is worthwhile. Evidence has been found in this study that older children can utilize the above-mentioned underlying rules in learning written Chinese better than the younger children, what is still unknown is how, and why and when.

10.3 A theory of instruction

10.3.1 What is taught and what is not: implicit and explicit knowledge in learning written Chinese

In Hong Kong, children learn to recognize characters at preschool level when they are expected to learn one to two characters a day, and no more than six to seven everyday in Grade One. Teachers do not differentiate simple characters and compound characters, and confine their teaching methodologies to those which are suitable to teach simple characters only. Simple characters are usually introduced to beginning readers first in order to demonstrate the basic strokes of the characters. These characters are initially presented with picture cues so that children can understand their meaning. Teachers emphasize the sound-symbol relationship between the syllable and the whole character, and pronunciation of these characters is achieved by rote memory.

Children learn to write the characters following a prescribed sequence of strokes, form and orientation. If we turn to the teaching guidelines of written Chinese in primary schools, very specific instructions are given in making the strokes correct, such as ‘the slanting stroke has a tiny hook at the end’, or ‘one vertical stroke has to be slightly longer than the other one’ in writing a character. Again it is assumed that the characters are learnt by memorizing the sequence of strokes one after the other. It may be true that young children have to learn some simple characters which have

no access to meaning and pronunciation systematically by rote memory, but children do not learn all the characters in this way.

Apparently, what is explicitly taught about the written Chinese is only applicable to the simple words which consist of strokes. No distinction is made in the orthographic structure of simple words and compound words, and the functional representation in the semantic radicals and phonological components have been largely overlooked. If we look at the textbooks for primary school children, only strokes are mentioned. In a writing exercise attached in Appendix 9, where children are required to revise twenty characters, individual strokes are laboriously repeated for every character. However, the common stroke-pattern *mouth* < 口 >, which appears in five characters on that page, all with different sizes and at various locations is not highlighted. It appears that the ability to locate the stroke-patterns, which has been found to be an important skill in the reading acquisition of Chinese, has been neglected by teachers. Furthermore, there is no mention of the functions of the stroke-patterns in formal teaching. Occasionally, the meaning of some semantic radicals, especially those which cannot be used as a character on their own, are being introduced, but they are taught in a piecemeal rather than systematic fashion. The function of the phonological components is seldom discussed, as most teachers believe that the relationship between the pronunciation of the phonological components and the whole character is rather weak.

The failure to differentiate the simple words and compound words in teaching results in a huge discrepancy between what is taught and what need to be learnt. This study has provided evidence that children have, in fact, learned more than memorizing individual characters according to the order of strokes. The following is the list of implicit knowledge that may be necessary for children in the process of learning to read.

1. Knowledge about the orthographic structure of written Chinese

- There is another level of orthographic unit - the stroke-patterns.
- There is a fixed set of stroke-patterns and they are the building blocks in compound characters.
- The same stroke-pattern can appear in different sizes and in different parts of the characters.

2. Knowledge about the function of the semantic radicals

- There is no pictorial association in most of the characters. The meaning of most of the characters is represented by the semantic radical.
- There is a relationship between the positions and functions of the semantic radicals.

3. Knowledge about the function of the phonological components

- There is a relationship between the pronunciation of the phonological component and the pronunciation of the whole character. The two possible ways of pronouncing words via the phonological components are by derivation or by analogy.
- Pronunciation can never be found in the semantic radicals.

10.3.2 Implications for teaching

No research has yet investigated if these units of implicit knowledge can be taught explicitly, and how far children can benefit from explicit teaching of these underlying rules of written Chinese. There is evidence in this study that some children fail to make a distinction between the orthographic structure of simple words and compound words, and they disregard the functions of the stroke-patterns. They are not aware that semantic radicals can never be pronounced, perhaps because

no one has explained to them the different functions of the semantic radicals and phonological components in ideophonic compound words. For the same reason, they cannot utilize the phonological components for pronunciation when they are faced with a novel word.

The failure to pay attention to the intra-character information probably creates difficulties for children learning to read Chinese. Some teachers have devised language games to highlight the importance of locating the semantic radicals and the phonological components. For example, children are asked to write down as quickly as possible characters which share the same semantic radicals or the phonological components. These games, which are invented by teachers themselves, can be adopted to design intervention studies to help poor readers to acquire written Chinese in a more systematic way.

To conclude, children do not learn to read Chinese by rote, and the learning process requires an analytical understanding of the underlying rules of written Chinese. It is hoped that the findings from this study will open up an entirely different approach in Chinese language teaching. Furthermore, it is stressed that what has been taught may not be the way children learn to read. It is only by investigating how children learn to read that teachers can acquire a true picture of how they understand the Chinese written system.

LIST OF APPENDICES

1. PSEUDOWORD PAIRS FOR THE PILOT STUDY IN STUDY 1	252
2. PSEUDOWORD PAIRS FOR THE MAIN STUDY IN STUDY 1	253
3. MEAN SCORES IN SAME AND DIFFERENT PARIS IN STUDY 1	254
4. PSEUDOWORDS AND NONWORDS FOR PILOT STUDY 2	255
5. PSEUDOWORDS AND NONWORDS FOR MAIN STUDY 2	255
6. MEAN SCORES OF PSEUDOWORDS AND NONWORDS IN STUDY 2	256
7. MEAN SCORES OF CORRECT PRONUNCIATION OF PSEUDOWORDS IN STUDY 4	257
8. MEAN SCORES OF CORRECT SEMANTIC CHOICE IN STUDY 4	257
9. EXTRACT FROM A WRITING EXERCISE IN PRIMARY 1 CHINESE TEXTBOOK	258

APPENDIX 1

1. 仿 仿	2. 杖 杖
3. 黔 黔	4. 燀 燀
5. 晃 晃	6. 愆 愆
7. 榘 榘	8. 蹠 蹠
9. 畔 畔	10. 劬 劬
11. 鯉 鯉	12. 彥 彥
13. 辛 昌	14. 劫 劫
15. 籍 能	16. 慷 慷
17. 迤 荀	18. 陂 瘥
19. 酹 贊	20. 斲 蠟

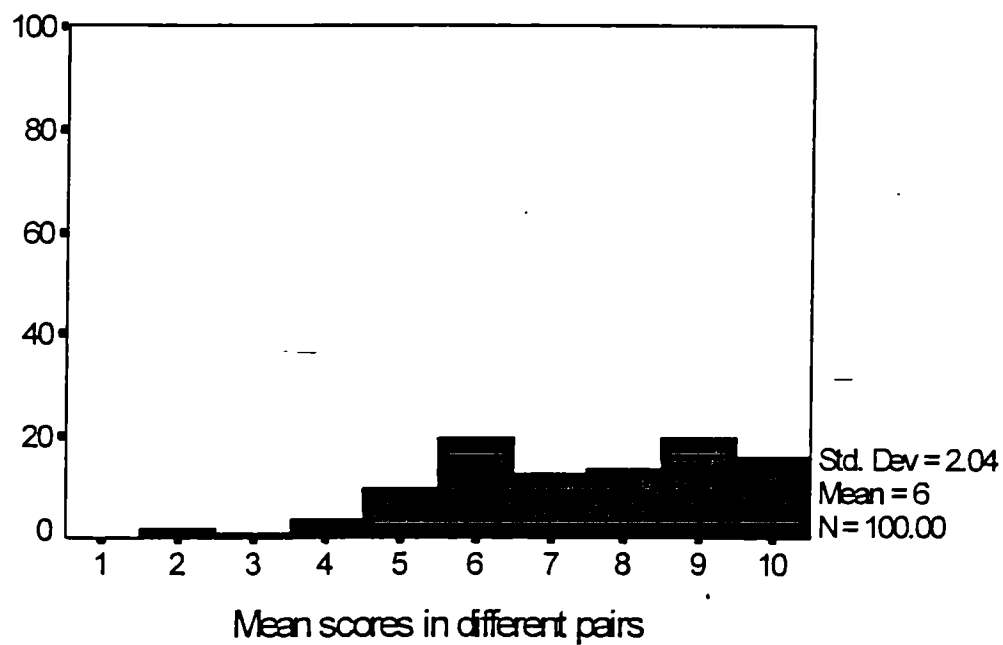
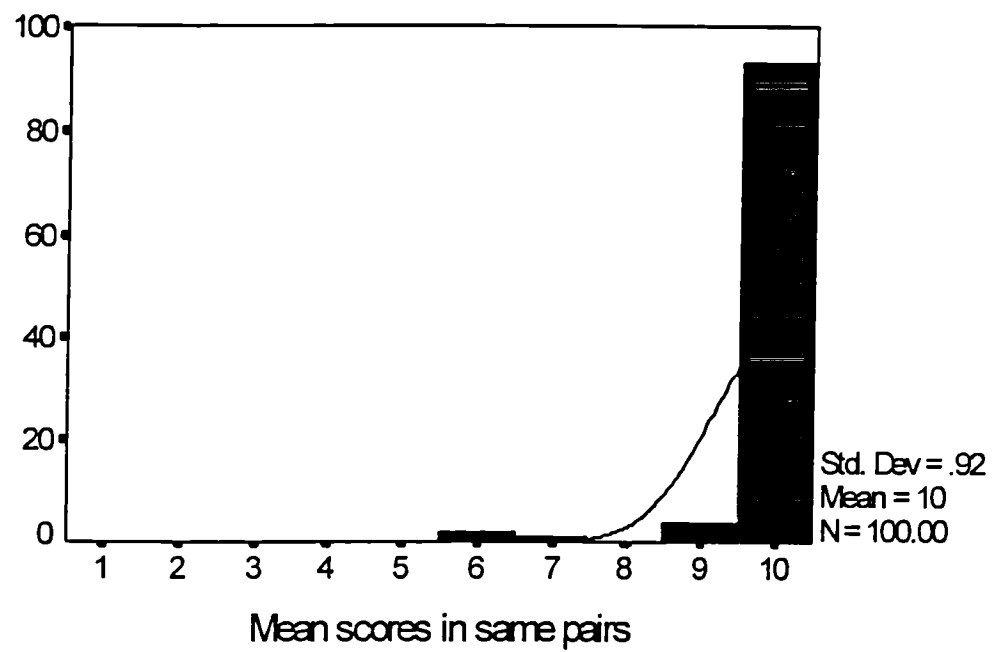
PSEUDOWORD PAIRS FOR THE PILOT STUDY IN STUDY 1

APPENDIX 2

1. 佻 佻	2. 标 标
3. 璽 璽	4. 稿 稿
5. 饒 饒	6. 折 折
7. 庑 庑	8. 髀 髀
9. 莖 莖	10. 蛸 蛸
11. 靚 靚	12. 鞣 鞣
13. 詆 羝	14. 魚 𩺰
15. 霽 霽	16. 瘰 瘰
17. 翅 翅	18. 釗 釗
19. 跣 跣	20. 槻 槻

PSEUDOWORD PAIRS FOR THE MAIN STUDY IN STUDY 1

APPENDIX 3



MEAN SCORES IN SAME AND DIFFERENT PAIRS IN STUDY 1

APPENDIX 4

PSEUDOWORDS

1. 𠄎	2. 痛	3. 琤	4. 洙	5. 搨	6. 息	7. 耐	8. 室	9. 咎	10. 倆
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

NONWORDS

1. 𠄎	2. 𠄎	3. 𠄎	4. 𠄎	5. 𠄎	6. 𠄎	7. 𠄎	8. 𠄎	9. 𠄎	10. 𠄎
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

PSEUDOWORDS AND NONWORDS FOR PILOT STUDY 2

APPENDIX 5

PSEUDOWORDS

1. 耐	2. 痛	3. 𠄎	4. 洙	5. 倆	6. 室	7. 琤	8. 搨	9. 咎	10. 息
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

NONWORDS WITH ILLEGAL GRAPHICS

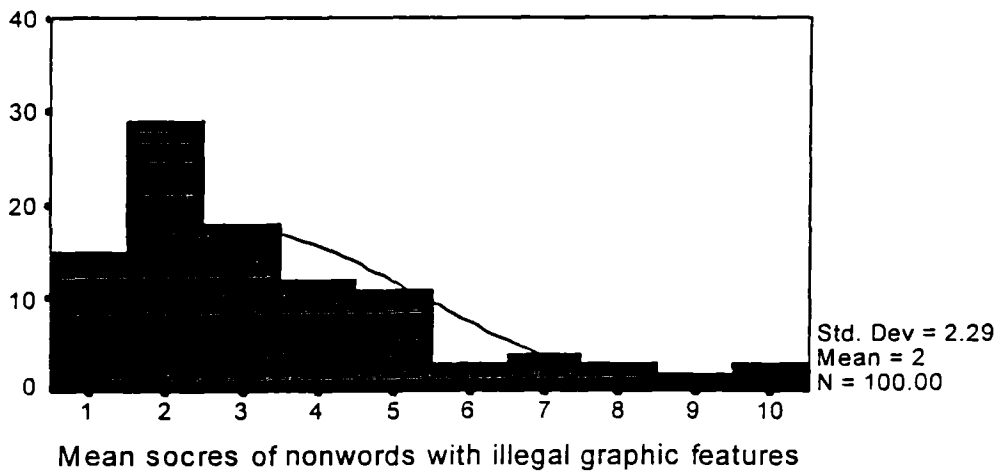
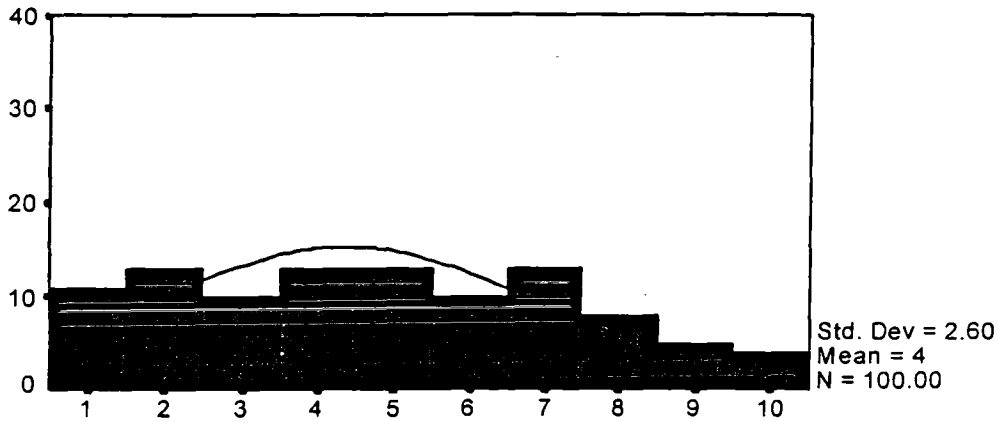
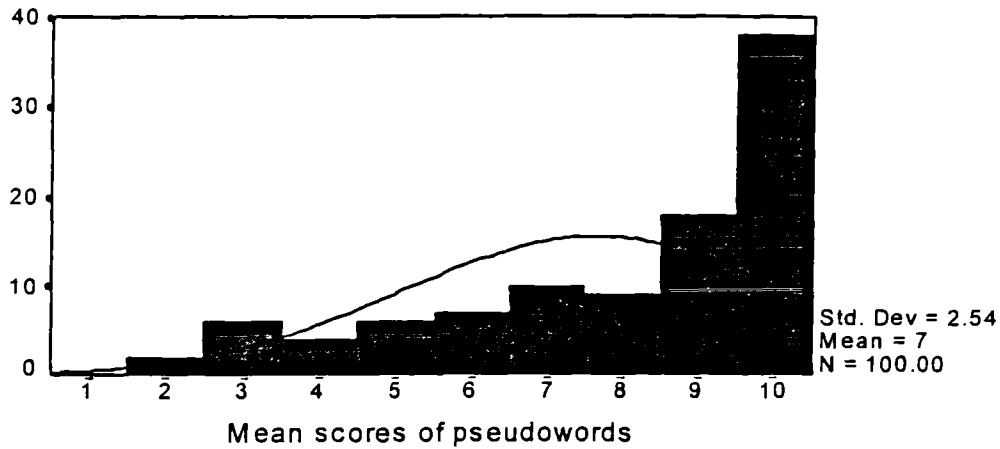
1. 𠄎	2. 𠄎	3. 𠄎	4. 𠄎	5. 𠄎	6. 𠄎	7. 𠄎	8. 𠄎	9. 𠄎	10. 𠄎
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

NONWORDS WITH ILLEGAL POSITIONS

1. 𠄎	2. 𠄎	3. 𠄎	4. 𠄎	5. 𠄎	6. 𠄎	7. 𠄎	8. 𠄎	9. 𠄎	10. 𠄎
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

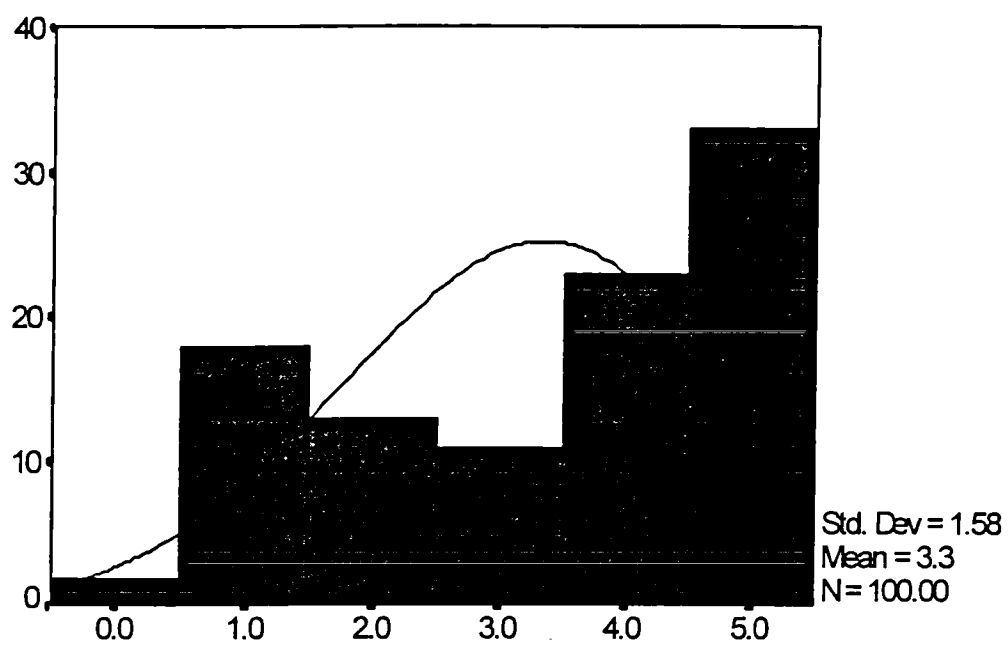
PSEUDOWORDS AND NONWORDS FOR MAIN STUDY 2

APPENDIX 6



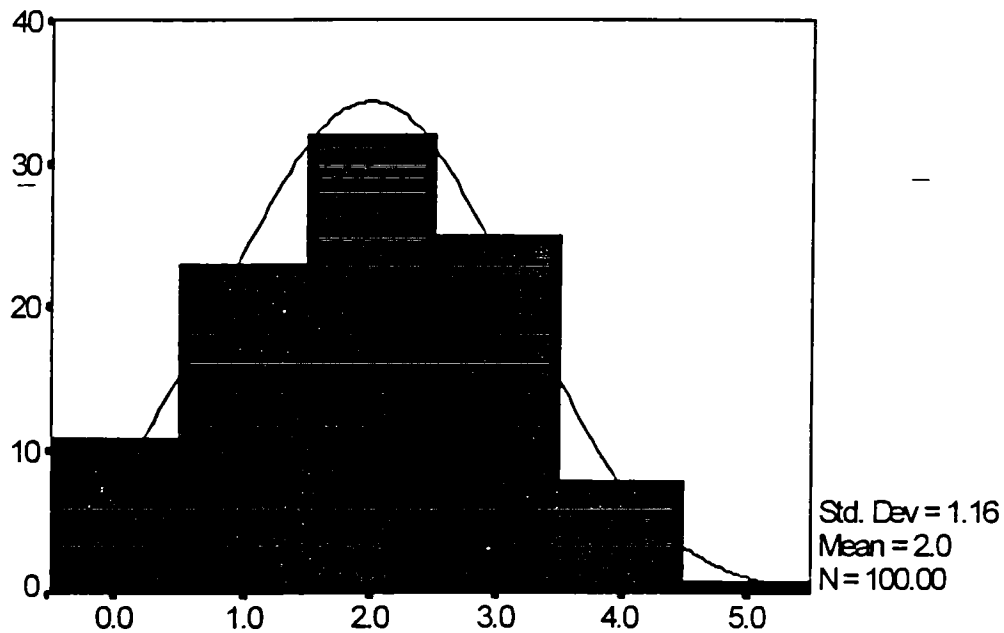
MEAN SCORES OF PSEUDOWORDS AND NONWORDS IN STUDY 2

APPENDIX 7

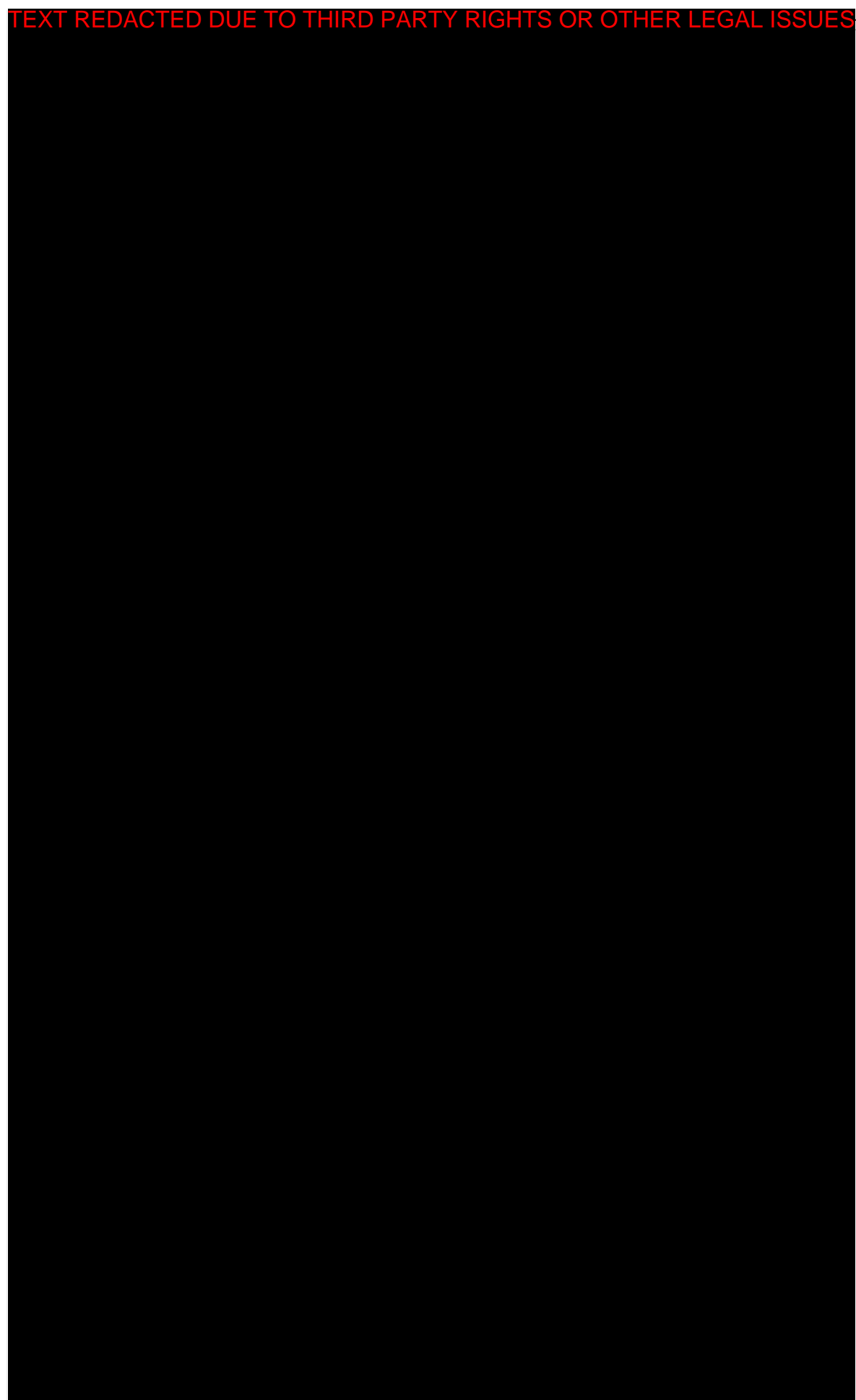


Mean scores of correct pronunciation of pseudowords in Study 4

APPENDIX 8



Mean scores of correct semantic choice in Study 4



REFERENCES

- Barron, R. W., & Baron, J. (1977). How children get meaning from printed words. Child Development, 48, 587-594.
- Beijing Foreign Language Press.(1989). Chinese characters. Beijing, Republic of China.
- Bryant, P., & Bradley, L. (1980). Why children sometimes write words which they do not read. In U. Frith (Eds.), Cognitive Processes in Spelling. London: Academic Press.
- Bryant, P., & Bradley, L. (1985). Children's reading problems : Psychology and Education. Oxford: Basil Blackwell Ltd.
- Chan, L. (1990). Preschool children's understanding of Chinese writing. Unpublished Master's thesis, Institute of Education, University of London.
- Chan, L., & Louie, L. (1992). Developmental trend of Chinese preschool children in drawing and writing. Journal of Research in Childhood Education, 6(2), 93-99.
- Chen, H. C., & Juola, J. F. (1982). Dimensions of lexical coding in Chinese and English. Memory and Cognition, 10, 216-224.
- Chen, M. J., Lau, L. L., & Yung, Y. F. (1993). Development of component skills in reading Chinese. In Sixth International Symposium on Cognitive Aspects of the Chinese Language. Taipei, Taiwan.
- Chen, M. J., Yung, Y. F., & Ng, T. W. (1988). The effect of context on perception of Chinese characters. Cognitive Aspects of the Chinese Language, 1, 27-39.
- Chen, Y.-P. (1993) Word recognition and reading in Chinese. Unpublished Ph.D. thesis, University of Oxford.

- Chen, Y.-P., & Allport, A. (1995). Attention and Lexical Decomposition in Chinese Word Recognition: Conjunctions of Form and Position Guide Selective Attention. Visual Cognition, 2(2/3), 235-268.
- Chen, Y. -P., Allport, D. A., & Marshall, J. C. (in press). What are the functional orthographic units in Chinese word recognition: the stroke or the stroke-pattern? Quarterly Journal of Experimental Psychology.
- Cheng, C. M. (1992). Lexical access in Chinese: Evidence from automatic activation of phonological information. In H. C. Chen & O.J.L. Tzeng (Eds.), Language processing in Chinese. Amsterdam: Elsevier Science Publishers.
- Cheng, C.-M. (1995). The acquisition of general lexical knowledge of Chinese characters in school children. In Seventh International Conference on The Cognitive Processing of Chinese and other Asian Languages. Hong Kong.
- Cohen, G. (1980). Reading and searching for spelling errors. In U. Frith (Eds.), Cognitive Processes in Spelling. London: Academic Press.
- Coltheart, M. (1978). Lexical access in simple reading task. In G. Underwood (Eds.), Strategies of information processing. London: Academic Press.
- Dai, B. Y., & Lu, J.P., (1985). Reading reform in Chinese primary schools. Prospects: Quarterly Review of Education, 15(1), 103-110.
- Dodd, B. (1980). The spelling abilities of profoundly pre-lingually deaf children. In U. Frith (Eds.), Cognitive Process in Spelling. London: Academic Press.
- Fan, L., Tong.L., & Song, J. (1987). The characteristics of Chinese language and children's learning to read and write. In D. Wagner (Eds.), The future of literacy in a changing world. Oxford: Pergamon Press.
- Fang, S.-P., Horng, R.-Y., & Tzeng, O. J. L. (1986). Consistency effects in the Chinese character and pseudo-character naming tasks. In H. S. R. Kao & R. Hoosain (Eds.), Linguistics, psychology, and the Chinese Language. Hong Kong: Centre of Asian Studies, University of Hong Kong.

- Fang, S. P., & Wu, P. (1989). Illusory Conjunctions in the perception of Chinese characters. Journal of Experimental Psychology: Human Perception and Performance, 15(3), 434-447.
- Ferreiro, E. (1978). What is written in a written sentence: A developmental answer. Journal of Education, 160, 25-39.
- Ferreiro, E. (1984). The underlying logic of literacy development. In H. Goelman, A. Oberg, & F. Smith (Eds.), Awakening to Literacy. London: Heinemann Education Books.
- Ferreiro, E. (1985). The relationship between oral and written language: the children's viewpoints. In M. M. Clark (Eds.), New Directions in the Study of Reading. London: The Falmer Press.
- Ferreiro, E., & Teberosky, A. (1982). Literacy before schooling. Exeter: Heinemann Educational Books.
- Flores d'Arcais, G. B. (1992). Graphemic, phonological and semantic activation processes during the recognition of Chinese characters. In H. C. Chen & O. J. L. Tzeng (Eds.), Language Processing in Chinese. Amsterdam: Elsevier Science Publishers.
- Flores d'Arcais, G. B., Saito, H., & Kawakami, M. (1995). Phonological and semantic activation in reading Kanji characters. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(1), 34-42.
- Frith, U. (1985). Unexpected spelling problems. In U. Frith (Eds.), Cognitive Processes in Spelling. London: Academic Press.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In M. Coltheart & J. Marshall (Eds.), Surface Dyslexia. London: Lawrence Erlbaum Associates.
- Gibson, E., & Levin, H. (1975). Psychology of Reading. Cambridge: M.I.T. Press.

- Glushko, R.J. (1979). The organization and the activation of orthographic knowledge in reading aloud. Journal of Experimental Psychology: Human Perception and Performance, 5, 674-691.
- Goelman, H., Oberg, A., & Smith, F. (1984). Awakening to literacy. London: Heinemann Educational Books.
- Goodman, Y. (1984). The development of initial literacy. In H. Goelman, A. Oberg, & F. Smith (Eds.), Awakening to Literacy. London: Heinemann Educational Books.
- Goodman, Y. M. (1986). Children coming to know literacy. In W. H. Teale & E. Sulzby (Eds.), Emergent Literacy, Writing and Reading Ablex.
- Goswami, U., & Bryant, P. (1990). Phonological Skills and Learning to Read. East Sussex: Lawrence Erlbaum.
- Henderson, L. (1982). Orthography and word recognition in reading. London: Academic Press.
- Henderson, L., & Chard, J. (1980). The Reader's Implicit Knowledge of Orthographic Structure. In U. Frith (Eds.), Cognitive Processes in Spelling. London: Academic Press.
- Ho, C. S., -H. (1989) The relevance of visual and phonological abilities for Chinese beginning readers. Unpublished Master thesis, University of Hong Kong.
- Ho, S.-H. (1994) A cross-cultural study of the precursors of reading. Unpublished Ph.D. thesis, University of Oxford.
- Hoosain, R. (1991). Psycholinguistic implications for linguistic relativity: a case study of Chinese. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Hu, & Huang (1990). The comprehensive knowledge of Chinese words. Wu Han: Hua Zhong Normal University Press.

- Huang, C. Y., & Hanley, J. R. (1992). The relationship between phonological awareness, visual skills and reading ability in the U.K., Hong Kong and Taiwan. In 25th International Congress of Psychology. Brussels, Belgium.
- Huang, H. S., & Hanley, J. R. (1993). Phonological awareness, visual skill and Chinese reading acquisition in first graders: A longitudinal study in Taiwan. In Sixth International Symposium on Cognitive Aspects of the Chinese Language. Taipei, Taiwan.
- Huang, T. J., & M.Y., W. (1992). From unit to Gestalt: Perceptual dynamics in recognising Chinese characters. In H. C. Chen & O. J. L. Tzeng (Eds.), Language Processing in Chinese. Amsterdam: Elsevier Science Publishers.
- Hue, C. W. (1992). Recognition Processes in Chinese Naming. In H.C.Chen & O.J.L. Tzeng (Eds.), Language processing in Chinese. Amsterdam: Elsevier Science Publishers.
- Hue, C. W., & Erickson, J. R. (1988). Short term memory for Chinese characters and radicals. Memory and Cognition, 16, 196-205.
- Jackendoff, R. (1993). Patterns in the Mind, Language and Human Nature. Hemel Hempstead: Harvester Wheatsheaf.
- Kamberelis, G. (1992). Markers of cognitive change during the transition to conventional literacy. Reading and Writing, 4(Dec.), 365-402.
- Karmiloff-Smith, A. (1992). Beyond modularity. A developmental perspective on Cognitive Science. Cambridge (MA): The MIT Press.
- Lai, C., & Huang, J. T. (1988). Component migration in Chinese characters: Effects of priming and context on illusory conjunction. In I. M. Liu, H. C. Chen, & M. J. Chen (Eds.), Cognitive aspects of the Chinese language. Hong Kong: Asian Research Service.
- Lavine, L. O. (1977). Differentiation of letterlike forms in pre-reading children. Developmental Psychology, 13, 89-94.

- Lee, L.-J. (1989) Emergent literacy in Chinese: print awareness of young children in Taiwan. Unpublished Ph.D. thesis, The University of Arizona.
- Lee, S.-y., Stigler, J. W., & Stevenson, H. W. (1986). Beginning reading in Chinese and English. In B. R. Foorman & A. W. Siegel (Eds.), Acquisition of reading skills: cultural constraints and cognitive universals. Hillsdale: Lawrence Erlbaum Associates.
- Liu, I. M., Chuang, C. J., & Wang, S. C. (. (1975). Frequency count of 40,000 Chinese words. Taipei: Lucky Books.
- Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1980). The development of strategies in spelling. In U. Frith (Eds.), Cognitive Processes in Spelling. London: Academic Press.
- Martin, S. E. (1972). Nonalphabetic Writing Systems: Some observations. In J. F. Kavanagh & I. G. Mattingly (Eds.), Language by Ear and by Eye: the relationships between Speech and Reading. Cambridge, Mass.: The MIT Press.
- McClelland, J.L., & Rumelhart, D.E. (1981) An interactive activation model of context effect in letter perception: Part I. An account of basic findings. Psychological Review, 88, 375-407
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic codes in visual word recognition. Memory and Cognition, 2, 309-321.
- Nunes Carraher, T., & Rego, L. L. B. (1982). Understanding the alphabetic system. In D. Rogers & J. A. Slododa (Eds.), The Acquisition of Symbolic Skills. NY: Plenum Press.
- Nunes, T., Bryant, P., & Bindman, M. (Submitted for publication). Morphological spelling strategies: Developmental stages and processes. Developmental Psychology.

- Peng, D. L., Li, Y. P., & Huang, F. M. (1995). Orthographic information in the identification of Chinese characters. In Seventh International Conference on The Cognitive Processing of Chinese and other Asian Languages. Hong Kong.
- Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. Journal of Memory and Language, 27, 59-70.
- Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. Journal of Experimental Psychology: Learning, Memory and Cognition, 17, 633-643.
- Perfetti, C. A., & Sulan, Z. (1995). Very early phonological activation in Chinese Reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(1), 24-33.
- Read, C. (1971). Pre-school children's knowledge of English phonology. Educational Review, 41, 1-34.
- Rozin, P., Poritsky, S., & Soetky, R. (1971). American children with reading problems can easily learn to read English represented by Chinese characters. Science, 171, 1264-1267.
- Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. Cognition, 19, 1-30.
- So, D., & Siegel, L. S. (1992). Learning to read Chinese: Semantic, syntactic, phonological and short-term memory skills in normally achieving and poor Chinese readers. Ontario. Unpublished manuscript.
- Stevenson, H. W., Stigler, J. W., Lee, S. Y., Lucker, G. W., Kimura, S. R., & Hsu, C.C. (1985). Cognitive performance and academic achievement of Japanese, Chinese and American children. Child Development, 56, 718-734.

- Stevenson, H. W., Stigler, J. W., Lucker, G. W., & Lee, S. Y. (1984). Reading disabilities: the case of Chinese, Japanese and English. Child Development, 53, 1164-1181.
- Taylor, I., & Taylor, M. M. (1983). The Psychology of Reading. New York: Academic Press.
- Temple, C., Nathan, R., Burris, N., & Temple, F. (1988). The beginning of Writing. Allyn and Bacon, Inc.
- Tolchinsky-Landsman, L. (1988). Form and meaning in the development of writing. European Journal of Psychology of Education, 3(4), 385-398.
- Tolchinsky-Landsman, L., & Levin, I. (1985). Writing in preschoolers: an age-related analysis. Applied Psycholinguistics, 6, 310-339.
- Tolchinsky-Landsman, L., & Levin, I. (1987). Writing in four-to six-year-olds: representation of semantic and phonetic similarities and differences. Journal of Child Language, 14, 127-144.
- Treiman, R. (1993). Beginning to Spell. New York: Oxford University Press.
- Tzeng, O. J. L., & Hung, D. L. (1980). Reading in a non-alphabetic writing system. In J. F. Kavangh & R. L. Veneky (Eds.), Orthography, Reading and Dyslexia. Baltimore: University Park Press.
- Tzeng, O. J. L., & Hung, D. L. (1981). Linguistic determinism: a written language perspective. In O. J. L. Tzeng & Singer (Eds.), Perception of print: reading research in experimental psychology. Hillsdale, NJ: Erlbaum.
- Wang, W. S.-Y. (1981). Language structure and optimal orthography. In O. J. L. Tzeng & H. Singer (Eds.), Perception of Print: Reading Research in Experimental Psychology. New Jersey: Lawrence Erlbaum Associates.
- Wang, W. S. Y. (1973). The Chinese Language. Scientific American, 228, 51-60.

- Yang, K.-T. (1990) A study of written language in Chinese first-grade children in Taiwan, the Republic of China. Unpublished Ph.D. thesis, Northwestern University.
- Yin, W.G., & Butterworth, B. (1992) Deep and surface Dyslexic in Chinese. In H.C.Chen & O.J.L. Tzeng (Eds.), Language processing in Chinese. Amsterdam: Elsevier Science Publishers.
- Zhang, G., & Simon, H. (1985). STM capacity for Chinese words and idioms: chunking and acoustical loop hypothesis. Memory and Cognition, 13, 193-201.
- Zhang, J. J., Zhang, H. C., & Peng, D. L. (1990). The recovery of meaning of Chinese characters in the classifying process. Acta Psychologica Sinica, 139-144.
- Zhou, Y. K. (1980). Precise Guide to Pronunciation with Chinese with Chinese Phonological Roots. Jilin People's Publishing Co.
- Zhu, Y. P. (1987). Analysis of cueing functions of the phonetic in modern China. In East China Normal University. (In Chinese)

